QUANTUM CHEMICAL STRUCTURE-REACTIVITY STUDIES OF ANTICANCER DRUG MOLECULES AND AN INSIGHT INTO DRUG-DNA INTERACTION

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

By

PUBALEE SARMAH

Registration No: 018 of 2009

Department of Chemical Sciences
School of Science and Technology
Tezpur University
Napaam, Tezpur-784 028
Assam, INDIA

November 2010

Dedicated to....

My Parents

Declaration

I hereby declare that the thesis entitled "***Quantum Chemical Structure-Reactivity Studies of Anticancer Drug Molecules and an Insight into Drug-DNA interaction***" being submitted to the Department of Chemical Sciences, Tezpur University, is a record of original research work carried out by me. Any text, figures, results or designs that are not of own devising are appropriately referenced in order to give credit to the original author(s). All sources of assistance have been assigned due acknowledgement. I also declare that neither this work as a whole nor a part of it has been submitted to any other university or institute for any other degree, diploma or award.

Date: 29/11/2010

Place: Tezpur

Pubalee Sarmah

(*Pubalee Sarmah*)

TEZPUR UNIVERSITY

(A Central University established by an Act of Parliament)

NAPAAM, TEZPUR – 784028

DISTRICT: SONITPUR :: ASSAM:: INDIA

Ph: 03712 – 267004
03712 –267005
Fax: 03712 –267006
03712 –267005
e-mail: ramesh@tezu.ernet.in

Dr. Ramesh C. Deka
Professor
Department of Chemical Sciences
Tezpur University

This is to certify that the thesis entitled "***Quantum Chemical Structure-Reactivity Studies of Anticancer Drug Molecules and an Insight into Drug-DNA interaction***" submitted by Ms. Pubalee Sarmah for the degree of Doctor of Philosophy of Tezpur University, embodies the record of original investigation carried out by her under my supervision. She has been duly registered and the thesis presented is worthy of being considered for the Ph. D. Degree. This work has not been submitted for any degree of any other University.

Date: 29/11/2010

Place: Tezpur

P Deka

Signature of the Supervisor

Acknowledgement

When I passed M.Sc. and joined as a research scholar in the Department of Chemical Sciences, Tezpur University, I was afraid as I believe "Working on a Ph. D. and writing a thesis is certainly a formidable endeavour." However, the last five years have been very enjoyable and I am indebted to many people for helping me in making this thesis possible. It is a pleasure to convey my gratitude to them all in my humble acknowledgment. First and foremost, I would like to thank my supervisor, Prof. Ramesh C. Deka, for giving me the opportunity to carry out research in such an interesting field of Computational Chemistry, his outstanding openness to interesting new ideas, perpetual enthusiasm, and commitment for his students. Apart from many scientific lessons, he taught me to be critical, focused and patient. Without his guidance and persistent help this thesis would not have been possible.

Next, I wish to express my warm and sincere thanks to all the faculty members and staff of the Department of Chemical Sciences, Tezpur University for their help and encouragement. Particularly, my special thanks go to Dr. A. K. Phukan who supported me with valuable advice and suggestions.

I would like to acknowledge Tezpur University and Department of Science and Technology, New Delhi for their financial support.

Many thanks go to Mrs. Surobhi Deka, Department of Mathematical Sciences, Tezpur University for her help in understanding the basic statistics, which is fundamental of QSAR modelling. In absence of her help, I would probably not have learnt as much as I did.

I gratefully acknowledge Dr. S. K. Ray, Department of Molecular Biologoy and Biotechnology, Tezpur University for providing me valuable literatures and suggestions.

It is my pleasure to thank all the research scholar of the Department of Chemical Sciences, Tezpur University for their support in various ways

during my stay in the department. Especially, I want to thank my labmates, Ajantaba, Bulumoniba, Kalyanda, Paritoshda, Nabanita, Subhi and Kusum for their help and support.

I am very thankful to my friends, Sagar (IITB), Kuldeep (NEIST, Jorhat) and Anupam (NewYork Univ.) who provided me different references and information used in this thesis.

I should also mention that the entire team of computer centre, Tezpur University supported me a lot by their quick help in any computer related problem.

My time at Tezpur University was made enjoyable in large part due to many hostel friends that became a part of my life. I am grateful for time spent with my friends, Mandakini, Kanan, Padmakshi, Nibedita, and Manashiba in several memorable trips and parties.

Collective and individual acknowledgments are also owed to my colleagues Biswa, Debarati, Narayan, and Navalaxmi whose presence perpetually refreshing and helpful.

Where would I be without my family? My parents deserve special mention for their inseparable support and encouragement. I know how much you care about me and pray for me. And to my three sisters and brother-in-laws, thanks for your consistent trust and just about everything you did for me. I must give a special thank to Tunba, who actually insisted on getting into Tezpur University for pursuing my Ph.D. work under the guidance of Prof. Ramesh C. Deka. I am blessed with this great family. Without your love and blessing I would never have been able to bring this to a conclusion.

Finally, I would like to express special thanks to Amal, whose faithful support provided constant comfort and enabled me to carry on when things didn't look that bright.

Department of Chemical Sciences
Tezpur University
Date: 29/11/2010

Pubalee Sarmah

Pubalee Sarmah

QUANTUM CHEMICAL STRUCTURE-REACTIVITY STUDIES OF ANTICANCER DRUG MOLECULES AND AN INSIGHT INTO DRUG-DNA INTERACTION

ABSTRACT

The goal of the present thesis is to focus the applications of density functional methods, both from conceptual as well as computational viewpoint to correlate mainly the structure and activity of several anticancer drug molecules. Density functional theory (DFT) based reactivity descriptors, viz., chemical potential, global hardness, electrophilicy and Fukui functions are used to calculate the reactive nature and active sites of the drugs. DNA damage is the underlying cause of mutation leading to cancer. The electronic properties of DNA responsible for its damage include electron affinities of the nucleobases that have been investigated in the framework of DFT. The thesis mainly deals with platinum anticancer drugs. In particular, an attempt has been made to correlate the activity and property of cis-platinum complexes with their reactivity descriptors by QSAR (Quantitative Structure Activity Relationship) and QSPR (Quantitative Structure Property Relationship) analyses. Further, we have intended to study the binding mechanism of platinum drugs with DNA by considering a clinically used potential anticancer drug, AMD473. We have also investigated the structure-activity relationship of nucleoside analogues and designed some new compounds with higher anticancer activity. The thesis is organized as follows.

Chapter 1 deals with general introduction of DNA starting from the brief history of its structure elucidation to the different factors leading to its damage. DNA damage can cause mutation which is responsible for the development of cancer. We have briefly presented the classification of cancer and different types of genes involved in cancer. An overview of platinum antitumor chemistry is presented which starts with the history of cisplatin, the first widely used platinum drug for therapeutic purposes. Subsequently, the advantages and disadvantages of cisplatin are discussed.

Mechanism of action of platinum drugs on molecular and cellular level is described in detail. In this introductory chapter, we have also discussed the contribution of different computational methods into the field of drug design emphasizing the growing role of quantum-mechanical methods. It includes the fundamentals of DFT and details of basis sets and energy functionals. We have presented the global and local reactivity descriptors which have been extensively used in the present study to investigate the reactive nature and structure-activity analysis of platinum drugs. Along with the platinum complexes, the usefulness of these descriptors in the structure-activity analysis of organic anticancer agents, viz., nucleoside analogues have also been studied. This chapter gives a description of the general QSAR methodology employed in the present work. Further we have discussed the applications of the hybrid quantum mechanics/molecular mechanics (QM/MM) method in the study of drug-DNA interaction. The scopes of the present investigation are also described in this chapter.

Chapter 2 describes the influence of different basis sets and exchange-correlation functionals in the determination of electron affinities of nucleobases. The low energy electrons produced in ionizing radiation can induce strand breaks in DNA via dissociative electron attachment. By trapping these electrons, nucleobases form radical anions which then participate in different chemical reactions that can lead to genetic damage. Thus determination of electron affinities of nucleobases has played a crucial role in the field of radiation induced mutagenesis. We have used GAMESS program for optimization of neutral and anionic DNA/RNA single bases, Guanine (G), Adenine (A), Cytosine (C), Thymine (T) and Uracil (U). Influence of basis sets in the study of adiabatic and vertical electron affinities of these biomolecules has been carried out by using hybrid exchange-correlation density functional B3LYP in connection with 6-31G, TZVP and 6-311++G** basis sets. Then with the better predicted basis set (6-311++G**), effect of some energy functionals namely, PBEOP, PBELYP and PBEVWN on electron affinity values of the nucleobases is investigated. Vertical electron affinities of all the systems calculated with B3LYP and PBEOP functionals are in agreement with experimental results. The adiabatic electron affinities of uracil and thymine are positive while using diffuse basis set and B3LYP functional. At all level of theories we found negative value of adiabatic electron affinities for adenine, cytosine and guanine. The higher value of adiabatic

electron affinities of guanine obtained at B3LYP/6-311++G** level confirms the formation of mixed dipole-covalent bound anion.

Chapter 3 presents DFT based reactivity analysis of some platinum(II) complexes including anticancer drugs, viz., cisplatin, carboplatin and oxaliplatin in gas and solvent phases. The global and local reactivity descriptors of the complexes, such as, hardness, chemical potential, electrophilicity index, Fukui function, and local philicity have been calculated using double numerical polarization (DNP) basis set in connection with three exchange-correlation functionals (BLYP, BOP and HCTH) using DMol^3 program. The reactivity trend of the complexes changes with the inclusion of solvent medium indicating oxaliplatin as the most reactive complex among them, in agreement with experimental reports. The local reactivity descriptors, Fukui function (f^+) and philicity (ω^+), maximum values of which represent the site for nucleophilic attack of a system, have also been calculated for all atoms of the complexes. It is found that, in all the systems Pt sites are prone to nucleophilic attack with maximum values of $f^+$and ω^+. Further, we have performed QSAR analyses of the selected systems in both gas and solvent phases using the calculated reactivity parameters. The structure–activity analysis performed using solvent phase derived electrophilicity values showed a good correlation with the experimental cytotoxicity values of the complexes.

Chapter 4 discusses the QSAR/QSPR analyses of several cis-platinum complexes in both gas and solvent phases. Different QSAR parameters including reactivity descriptors have been calculated from DFT calculations of the complexes which are carried out at BLYP/ DNP level. Some molecular mechanics parameters have also been derived from the MM+ computations. From QSAR analyses of the complexes against A2780 human ovarian adenocarcinoma cell line and its cisplatin resistant subline (A2780Cp8), we have found that DFT derived reactivity descriptors, in particular electrophilicity and philicity in combination with energy of next LUMO can correlate drug activity of cis-platinum complexes remarkably in both gas and solvent phases. However, we have found that predictability of each model increases in solvent medium. In addition, we have calculated logarithimic n-octanol/water partition coefficient ($\log P_{o/w}$) values of 24 platinum complexes with different

leaving and carrier ligands, by noting its importance in drug action, metabolism and receptor binding. QSPR analyses of these complexes against 0% (extrapolated), 20%, 30%, 40% and 50% MeOH reveal that DFT based reactivity descriptors in combination with molar refractivity of carrier ligand and van der Waals surface area can be used for prediction of hydrophobicity of platinum complexes. The $\log P_{o/w}$ values of an additional set of 20 platinum complexes have also been modeled with the same descriptors. We have investigated predictive ability of the QSPR model by calculating $\log P_{o/w}$ of 4 compounds in the test set and found their predicted values to be in good agreement with the experimental values. For each QSPR model, we have found that statistical parameters of the models become more significant with the inclusion of solvent medium.

Chapter 5 is closely related to chapter 4. It describes the QSAR analysis of some carbocyclic analogues of nucleosides against murine leukemia cell line (L1210/0) and human T-lymphocyte cell lines (Molt4/C8 and CEM/0) in both gas and solvent media. Carbocyclic nucleosides are compounds in which the furan ring of the nucleoside is replaced by a carbocyclic system. This modification makes the molecules more resistant to hydrolases than the natural nucleosides. Among the various DFT and MM+ descriptor, we have found that energy of the next LUMO (E_{NL}), electrophilicity (ω) and van der Waals surface area (SA) are the main independent factors contributing to the anticancer activity of nucleoside analogues. QSAR equations with only two parameters for 14 carbocyclic nucleosides show good statistical quality both in regression ($r^2>0.90$) and LOO cross-validation ($r^2_{CV}>0.86$). After analyzing the parameters, we have designed 10 new compounds with rather high anticancer activities than those of the 14 compounds against the same cancer cell lines. The QSAR models developed for an additional set of 20 nucleoside analogues with three descriptors i.e., ω, E_{NL}, and SA provide significant statistical parameters in both gas and solvent media.

Chapter 6 includes detailed mechanistic study of a new promising anticancer agent, *cis*-$[PtCl_2(NH_3)$(2-picoline)], known as AMD473. Binding mechanism of platinum drug with DNA involves hydrolysis of the drug inside the cell before reaching DNA where both leaving ligands are replaced by aqua species. Hydrolysis of AMD473

along with its interaction with guanine DNA base has been investigated using *ab initio* Hartree–Fock (HF) and density functional levels of theory in gas phase and aqueous solution. Four different paths of hydrolysis have been studied by considering the replacement of two Cl atoms *trans* to NH_3 and 2-picoline ligands. We have used Gaussian 03 program for the optimization of all species involved in different reaction pathways. The nature of the stationary points located on the potential energy surface (PES) has been checked by vibrational analysis. The rate of hydrolysis of Cl atom *trans* to 2-picoline group is higher than that of Cl atom *cis* to 2-picoline group due to steric effect experienced by the axial picoline ligand. The structural analysis for the intermediates and transition states for guanine binding reactions showed that hydrogen bonds with the guanine O6 atom play an important role in stabilizing these species. All reaction pathways have been confirmed through the IRC analysis.

In **Chapter 7** we reports the four stabilities of different forms of AMD473-DNA adducts by discussing the energies as well as structural differences between them. Due to the asymmetric structure of AMD473 drug, it can form four stereoisomers with DNA. These are **(a)** 2-picoline *trans* to 3'-G and 2-methyl group directed to the DNA backbone, **(b)** 2-picoline *trans* to 3'-G and 2-methyl group directed to the DNA major groove center, **(c)** 2-picoline *trans* to 5'-G and 2-methyl group directed to the DNA backbone, and **(d)** 2-picoline *trans* to 5'-G and 2-methyl group directed to the DNA major groove center. We have used quantum mechanics/molecular mechanics (QM/MM) based two layer ONIOM (Our own N-layered Integrated molecular Orbital and molecular Mechanics) method, as implemented in the Gaussian 03 program to investigate the stabilities of these AMD473-DNA adducts. Further, we have studied the possibilities of proton transfer between DNA bases of the most stable drug-DNA adduct which leads to mutation of the DNA. From the calculations, it is found that adduct **b** is the most stable configuration among all the possible adducts and thus it is selected to study the proton transfer. From the proton transfer analysis we found that the single proton transfer between N1 position of guanine (G) and N3 position of cytosine (C) gives an energetically stable structure.

Chapter 8, the last chapter of the thesis summarizes the salient observations emerging out of the entire work and future scopes of the present investigation.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

BLYP	Becke exchange functional + Lee-Yang-Parr correlation functional
B3LYP	Becke's three parameter exchange functional + Lee-Yang-Parr correlation functional
BOP	Becke exchange functional + One-Parameter progressive correlation functional
COSMO	ConductOr-like Screening MOdel
DFT	Density Functional Theory
DNA	Deoxyribonucleic Nucleic Acid
DNP	Double Numerical with Polarization basis set
EA	Electron Affinity
ECP	Effective Core Potential
HCTH	Hamprecht, Cohen, Tozer and Handy exchange-correlation functional
HOMO	Highest Occupied Molecular Orbital
HPA	Hirshfeld Population Analysis
LUMO	Lowest Unoccupied Molecular Orbital
ONIOM	Our-own-N-layered Integrated Molecular Orbital + Molecular Mechanics
PCM	Polarizable Continuum Model

PBEOP	Perdew-Burke-Ernzerhof exchange + One-Parameter progressive (OP) correlation functionals
PBELYP	Perdew-Burke-Ernzerhof exchange + Lee, Yang, and Parr correlation functionals
PBEVWN	Perdew-Burke-Ernzerhof exchange + Vosko, Wilk, Nusair correlation functionals
QSAR	Quantitative Structure-Activity Relationship
QSPR	Quantitative Structure-Property Relationship
QM/MM	Quantum Mechanics/Molecular Mechanics
TZVP	Triple-Zeta Valence Polarized basis set
UFF	Universal Force Field

LIST OF FIGURES

LIST OF TABLES

Chapter 1

General Introduction

An overview about DNA and anticancer drug molecules are given, with special attention to the platinum complexes and nucleoside analogues. In this introductory chapter, we have also presented electronic structure methods used in computational chemistry including density functional theory (DFT).

1.1 An Overview of DNA

1.1.1 History

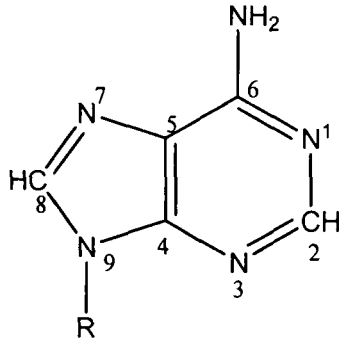
Deoxyribose nucleic acid (DNA) was first discovered by Friedrich Meischer in 1868; however, its role in heredity was not recognized before 1944 until Avery and co-workers reported that DNA and not the proteins were the carriers of genetic information.[1] In 1950 Erwin Chargaff reported that DNA composition is species specific i.e its amount varies from one species to another.[2] Chargaff also reported that amount of adenine equals to the amount of thymine and the amount of guanine equals to the amount of cytosine in DNA for every species.[3] The milestone in the DNA research was the discovery of its helical structure. Rosalind Frankllin, a young research associate in John Randalls lab at Kings College in London was the first scientist to discover that phosphate-ribose backbone lies on the outside of the DNA. She also elucidated the basic helical structure of DNA on the basis of X-ray crystallography technique. This discovery was the key step of the structural elucidation of DNA by Watson and Crick in 1953 by X-ray fiber diffraction technique.[4] They realized that the quantitative relationship between the nitrogeneous bases in DNA as suggested by Chargaff could be due to the complementary adenine-thymine and guanine-cytosine base pairing and on this basis they discovered the hydrogen bonding between these bases which is now known as Watson-Crick hydrogen bonding.

From this information Watson and Crick modelled a right handed antiparallel double helical structure of DNA where phosphate backbone lied outside the helix and the bases held through hydrogen bonding are pointed towards the centre of the helix. This discovery offers the concept of how information could pass from one generation to next by synthesis of DNA complementary strands from parent strands.

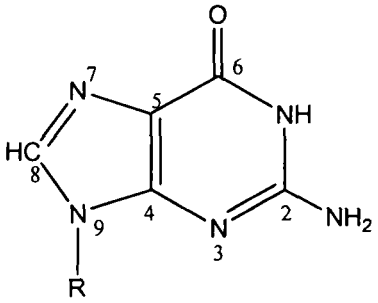
1.1.2 Structure of DNA

DNA is a polymer of deoxyribose nucleotides. The nucleotides consist of the components: the deoxyribose sugar ring; four nucleic acid bases; Adenine (A), Guanine (G), Thymine (T), Cytosine (C), and the phosphate linkers (Figure 1.1).

Purines

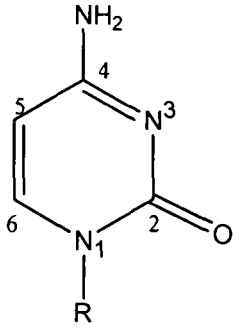


Adenine

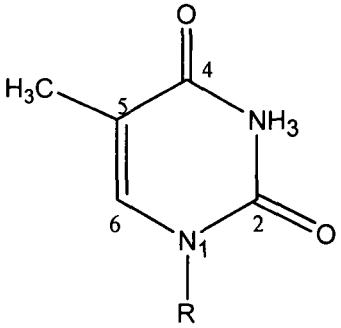


Guanine

Pyrimidines



Cytosine



Thymine

Deoxyribose sugar ring with phosphate group at 5' position

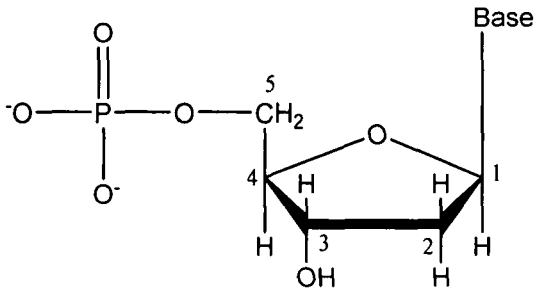


Figure 1.1: Schematic drawings of the nucleic acid bases and deoxyribose monophosphate.

DNA is linked by phosphate groups and hydrogen bonds between nucleotides of opposing strands forming a double helical structure as shown in Figure 1.2. In a nucleotide the base is joined to the C1 carbon of the sugar moiety and phosphate groups form bonds with either the 5'C or 3'C of the sugar. The two strands of the

DNA go in opposite direction and read as 5' to 3' direction. Chemically and biologically these ends are quite distinct. This property gives polarity to each DNA strand. The outer edges of strands are formed by nucleotides with altering sugar molecules and phosphate groups. The two strands are held together by hydrogen bonds between individual bases. In Watson-Crick base pairing the bases are hydrogen bonded with strict complementary base pairing according to the Chargaff rules.[3] Adenine on one strand pairs only with thymine on the other strand (A-T) and guanine pairs only with cytosine (G-C) as illustrated in the Figure 1.2.

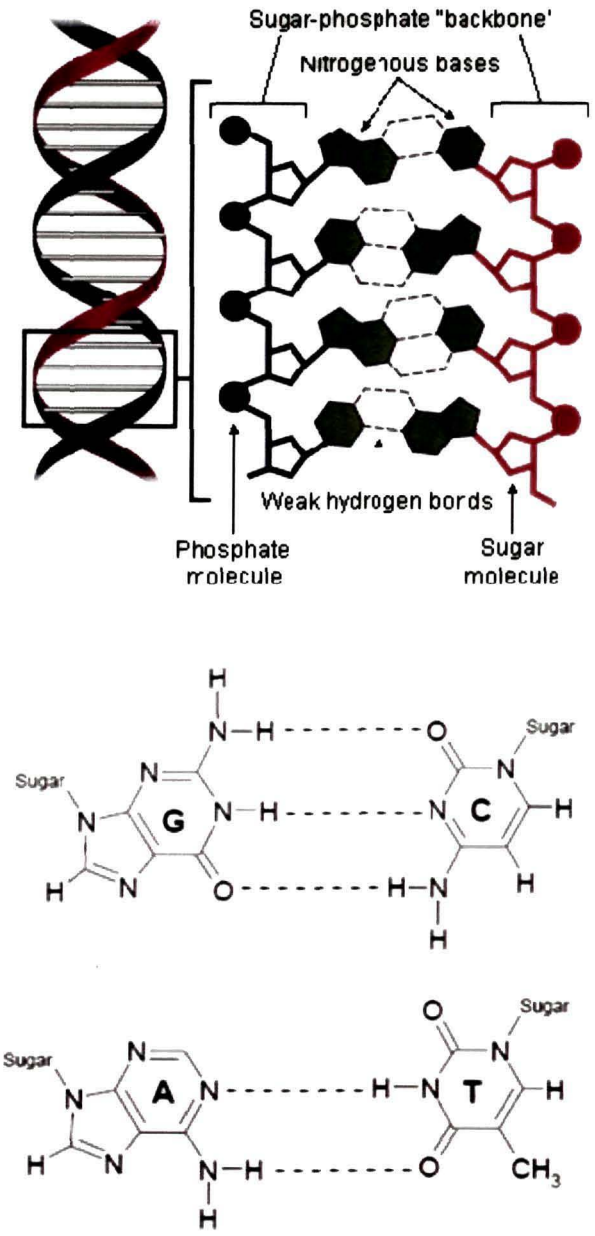


Figure 1.2: The DNA double helical structure and complementary base pairing.

The G-C pair has three hydrogen bonds while the A-T pair has only two and thus GC base pair is stronger than the AT. Several types of base pairing are possible, but the Watson-Crick base pairing is energetically favourable. The bases are stacked near the centre of the cylindrical helix providing considerable stability to the double helix.

The sugar moiety of DNA is one of the most flexible and dynamic part which can undergo various kinds of torsions. Thus sugar-phosphate backbone of DNA is rather flexible. Figure 1.3 represents the five endocyclic torsion angles ($\nu_{0\text{-}4}$) of pentose sugar geometry. The most common sugar pucker conformations are C2'-endo and C3'-endo as shown in Figure 1.3. These sugar puckering determines the shape of the helix, whether the helix will exist in the A form or in the B-form.

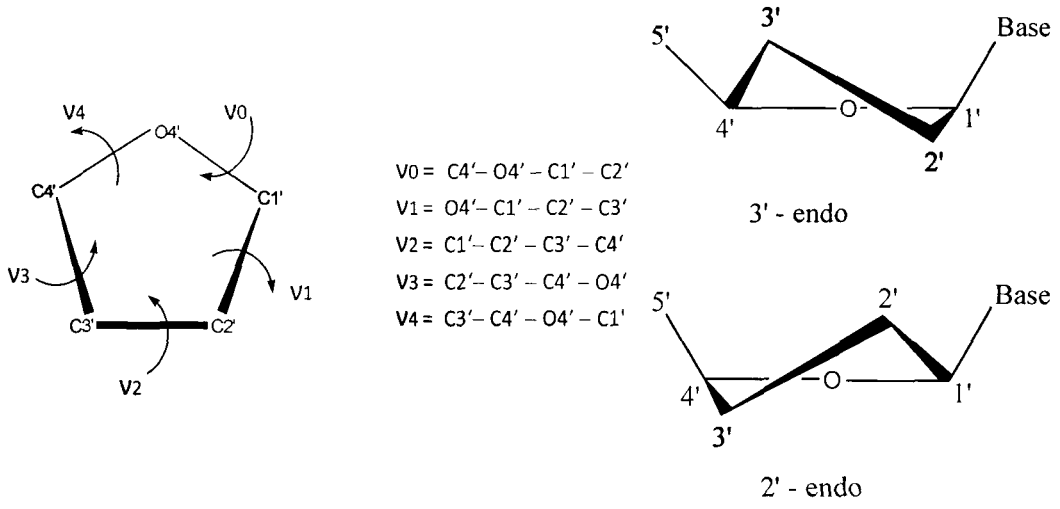


Figure 1.3: Torsion angles of pentose sugar in the phosphate backbone and C3′-endo and C2′-endo sugar puckering.

Figure 1.4 shows the most common conformations of DNA which are known B-, A-, and Z-DNA. The model proposed by Watson and Crick using X- ray crystallography technique is now known as B-DNA, which is the most frequently occurring conformation in nature. The ideal form of B-DNA is a right handed double helix with antiparallel backbone chains. The helix makes one complete turn approximately every 10.5 base pairs. The base pairs are perpendicular to the helix axis and twisted by ~36°. The distance between two neighbouring base pairs is 0.34 nm. The form of the ribose sugar is C2'-endo. The intertwined strands make two

distinct grooves, a wide major groove and a narrow minor groove. These two grooves provide very distinct surfaces with which proteins can interact. Many proteins interact in the space of the major groove, where they make sequence specific contacts with the bases. In addition, a few proteins are known which can interact *via* minor groove.

In a solution with higher salt concentrations, the DNA structure may change from B to A form, which is still right-handed, but makes a turn at every 2.3 nm and there are 11 base pairs per turn. In A-DNA the bases are much more tilted than B-DNA (~20°). The sugar pucker found in B-DNA changes from C2' endo to C3' endo in A-DNA.

One of the most dramatic conformations of DNA is Z-DNA, which is a left handed helix.[5] Like B-DNA, the two strands of Z-DNA are antiparallel and joined by Watson-Crick base pairing. In standard B-DNA the bases are usually in the *anti* conformation, which is sterically more favourable. In contrast to B-DNA, the bases in the Z-DNA helix alternate between the *anti* conformation and have unusual *syn* conformation. The dinucleotide repeat causes the backbone to follow a zigzag path, giving rise to the name Z-DNA. One turn spans 4.6 nm, comprising 12 base pairs. In Z-DNA, the major groove is disappeared into a nearly flat surface. The one visible groove is deep and narrow that corresponds to the minor groove of Z-DNA.

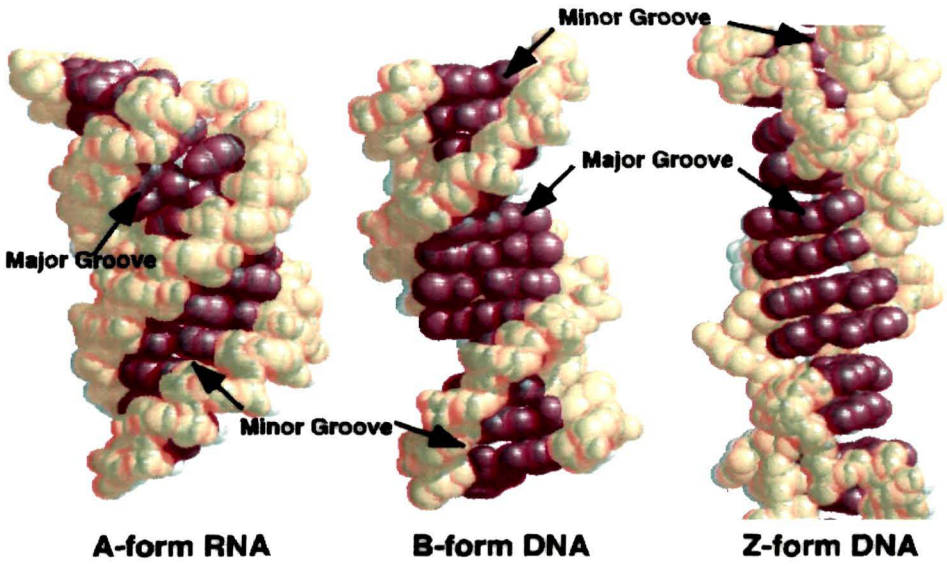


Figure 1.4: Schematic drawing of A-DNA, B-DNA, and Z-DNA.

1.1.3 DNA Damage

Since DNA is a chemical it is constantly damaged by chemical reactions. Chemical damage in DNA is called a lesion. Many environmental factors can damage this molecule. Chemical damage can be very deleterious to cells because the DNA may not be able to replicate over the damaged area and so the cells could not multiply. Even if the damage does not block replication, replicating over the damage can cause mutation, which is responsible for the development of cancer. There are many different ways that DNA can be changed, resulting in different types of mutations. Mutations are of two types; spontaneous mutations and induced mutations. The spontaneous mutation includes tautomerism, depurination, de-amination etc. The induced mutations at molecular level can be caused by chemicals and radiations.

The most common chemical induced mutation is the treatment of cells with deaminating agents. When adenine is deaminated the resulting hypoxanthine molecule pairs with cytosine during replication as shown in the Figure 1.5 and thus incorporating C instead of T at that position. In a subsequent replication, the C will pair with the correct G, causing an AT to GC transition in the DNA. DNA may

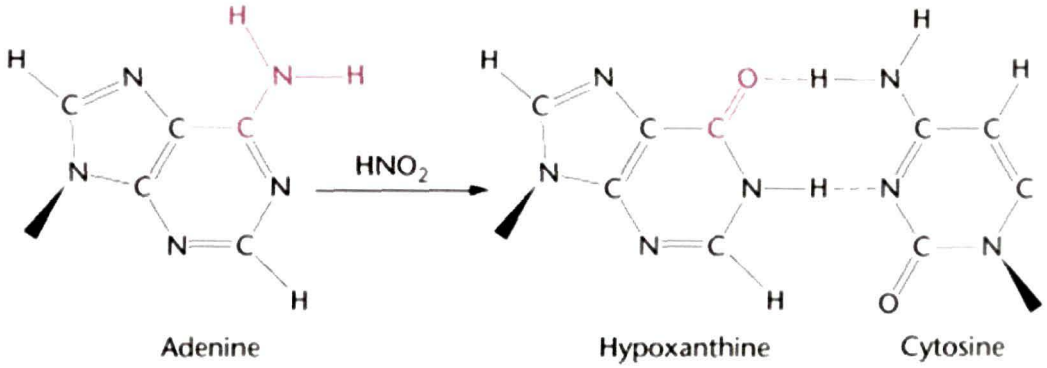


Figure 1.5: Modified base pairing created by deaminating agent, nitrous acid.

undergo oxidative damage by reactive oxygen that has more electrons than molecular oxygen. The base analogs are chemicals that can substitute for normal nucleobases in nucleic acids. Incorporation of a base analog can be mutagenic because the analog after pairing with the wrong base, leads to changes in the base pairing of DNA. Many chemicals can form interstrand cross-links in the DNA, in

which two bases in the opposite strands of the DNA are covalently joined to each other. DNA replication is blocked by cross-links which causes replication arrest and cell death.

The major source of radiation induced damage to DNA is UV irradiation due to sun exposure. Two nucleotide bases, cytosine and thymine of DNA are most vulnerable that can change their properties. UV light can induce adjacent thymine bases in a DNA strand to pair with each other, as a bulky dimer called pyridine dimer. On the other hand, when ionizing radiation interacts with living cell, several reactive species are formed among which most abundant species are electrons and hydroxyl radicals. However, the low energy electrons produced in significant amount during ionizing radiation have long been ignored as a potential DNA damaging factor. It has been believed that electrons produced in the cell exposed to ionizing radiation, unlike hydrogen or hydroxyl radicals, are inactive towards the biopolymer. Recently, experimental and theoretical studies have demonstrated that even at very low energies, electrons may induce strand breaks in DNA *via* dissociative electron attachment.[6,7] Numerous studies have shown that nucleobases provide trapping sites for these electrons. The resultant radical anions then participate in chemical reactions that can lead to the permanent alteration of the original bases and to genetic damage. Nucleic acid base anions thus play a central role in the electron-driven aspects of radiation-induced mutagenesis. In this context, the determination of electron affinities of DNA and RNA bases have significance in the study of radiation damage as well as excess electron transfer through DNA. Owing to the importance of electron affinities in DNA damage, we have investigated this property of nucleobases using various theoretical methods and presented in Chapter 2 of the thesis.

1.2 Cancer

Cancer is a very widespread disease, and is caused by cells which divide in an unregulated manner. In 2005 it accounted approximately 13% of all deaths.[8] Nowadays cancer is the second-leading cause of death in the Western world. It is a genetic disorder involving dynamic changes in the genome leading to uncontrolled growth of cells, which can affect and damage adjacent normal tissues. Cancer occurs after normal cells have been transformed into neoplastic cells through

alteration of their genetic material due to DNA damage and the abnormal expression of certain genes. Neoplastic cells usually exhibit chromosomal abnormalities and the loss of their differentiated properties. These changes lead to uncontrolled cell division and result in the invasion of previously unaffected organs, a process called metastasis.

Solid malignancies form lumps and liquid tumors circulate freely in the bloodstream. It can be caused or at least initiated by both external (carcinogens, tobacco, and radiation) and internal (hormonal effect, inherited mutations or immune deficiency) factors. Cancer can be broadly classified into four classes as (1) *Carcinomas*- characterized by cells that cover internal and external parts of the body such as lung, breast, colon etc. (2) *Sarcomas*- characterized by cells that are located in bone, fat, connective tissue, muscle etc. (3) *Lymphomas*- originated from lymphatic nodes and immune system tissues (4) *Leukemias*- originated from the bone marrow and often accumulate in the blood stream.

Generally several genes are anticipated to be involved in cancer developments. The two main types of genes that are playing role are oncogenes (growth promoting) and tumor suppressor genes (growth suppressing).[9] Oncogenes are related to normal genes called as 'proto-oncogenes' that normally control cell growth. The mutated or otherwise damaged versions of these genes are called oncogenes which can induce cancerous growth by instructing cells to synthesize cell growth and division stimulator proteins. There are some controlling tools for cell growth, namely, growth factors, receptors, signaling enzymes, and transcription factors. Growth factor binds to a cell-surface receptor triggering an intracellular signaling pathway and activate the transcription factor inside the cell. Consequently the activated transcription factors trigger the genes required for cell growth and division. When oncogenes are in control of cellular growth, they transform the growth-signalling pathway to be constantly active resulting uncontrolled cell growth.

Second group of genes responsible for cancer are the '*tumor suppressor genes*'. A tumor suppressor gene is a gene that reduces the probability of turning a cell in a multicellular organism to a tumor cell. Mutation or deletion of such gene will increase the probability of the formation of tumor. Tumor suppressor gene mutations have been found in many cancers. Most of these mutations are acquired, not inherited. Acquired mutations of the gene for *p53* appear in a wide range of cancers, including lung, colorectal, and breast cancer. The *p53* protein is involved in

the pathway of cell suicide called apoptosis. In cells that have undergone DNA damage, the *p53* protein acts like a "turn-off switch" halting cell division.

Another type of genes implicated in cancer are '*DNA repair genes*'. The normal function of these genes is to correct errors that arise when cells duplicate their DNA prior to cell division. Mutations in DNA repair genes can lead to a failure in repair, which in turn allow subsequent mutations to accumulate. Cancer may begin because of the accumulation of mutations involving oncogenes, tumor suppressor genes, and DNA repair genes.

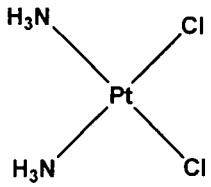
1.3 Treatment of Cancer

There are several classical approaches used to treat cancer, such as surgery, chemotherapy, radiation therapy, or combinations of these treatments. The choice of treatment is highly variable and dependent on a number of factors including the type, location and stage of disease and the health status of the patient.

The cases in which the cancer is detected at an early stage, surgery may be sufficient to cure the patient by removing all cancerous cells. The treatment of radiotherapy, however dependent on mechanistic selection of more cancer cells to damage than the normal cells. Surgery and radiation can be effective only if the cancer has not already metastasized because these treatments are local.[10] Advanced cancers whose cells have undergone many different mutations develop metastasis, causing 90% of cancer deaths. At this stage the disease is of systemic nature, requiring a systemic treatment and therefore often followed by chemotherapy. Chemotherapy is very important in treating blood cancers because often the tumour cells are dispersed all over the body and cannot be removed by surgery or reduced by radiotherapy. Furthermore, chemotherapy has huge potential for a number of diseases in spite of extensive spreading metastasis.[10] It uses drug molecules to treat cancer in a non specific way. The first chemotherapy treatment of cancer came after the observation of very low white blood cell counts in persons exposed to the chemical warfare agent mustard gas in the 1950s. Anticancer drugs taken orally, by injection, or intravenously are distributed throughout the body, and selectively kill the rapidly proliferating cancerous cells.

1.4 Platinum Anticancer Drugs

The first platinum complex to be used in the treatment of cancer was *cis*-diamminedichloroplatinum(II), clinically known as cisplatin (1, Figure 1.6), which was first synthesized by Michel Peyrone in 1845.[11]



1

Figure 1.6: Structural formula of cisplatin (**1**)

However, its activity against cancer remained unknown until 1964, when Rosenberg and his co-workers observed that platinum electrodes used in one of the experiments inhibit cell division.[12,13] The experiment was originally designed to measure effects of electrical current on bacterial growth of Escherichia coli cells. They observed filamentous growth of the cells, which normally occurs due to the block of DNA replication. Surprisingly, they found that the effect did not come from the induced electrical field. Rather, the compound, $Pt(NH_3)_2Cl_2$ formed by reaction of platinum from the electrodes with NH_4Cl in the buffer stopped cell division and induced filamentous growth in the bacteria. The drug entered clinical trials in 1971. Subsequently, a series of successful experiments on cancer cell lines with cisplatin were performed and approval by the Food and Drug Administration (FDA) of USA was granted in 1978. Cisplatin has since developed into one of the most widely used anticancer drugs and a paradigm for the treatment of testicular, head and neck, ovarian, esophageal, and non-small cell lung cancer.[14-16]

1.4.1 Mode of Action of Cisplatin

1.4.1.1 Hydrolysis

Cisplatin is usually administered intravenously rather than orally because of solubility problems. Since this intravenous protocol is associated with several side effects, researches have been carried out to find an alternative route of administration. Recently, controlled release drug delivery system for cisplatin has

been successfully developed by encapsulating the drug into different nano-scaled carriers.[17,18] In the bloodstream, high concentration of chloride ions (100 mM) suppresses the hydrolysis of cisplatin and maintains the complex in a neutral state. After reaching its target, cisplatin is taken up in the cell. The cellular uptake of cisplatin was classically believed to occur mainly by passive diffusion.[14,19] However, more and more evidences of involvement of active transport in the uptake are reported.[20-22] Recent studies have demonstrated that copper transporter Ctr1p mediates the influx and efflux of Pt-based antitumor agents both in yeast and in mammals.[23,24] Within the cell, the lower concentration of chloride ion (between 2 and 30 mM) facilitates the hydrolysis of cisplatin to yield chloro-aqua $[Pt(NH_3)_2Cl(H_2O)]^+$ and diaqua $[Pt(NH_3)_2(H_2O)_2]^{2+}$ species as shown in Figure 1.7.

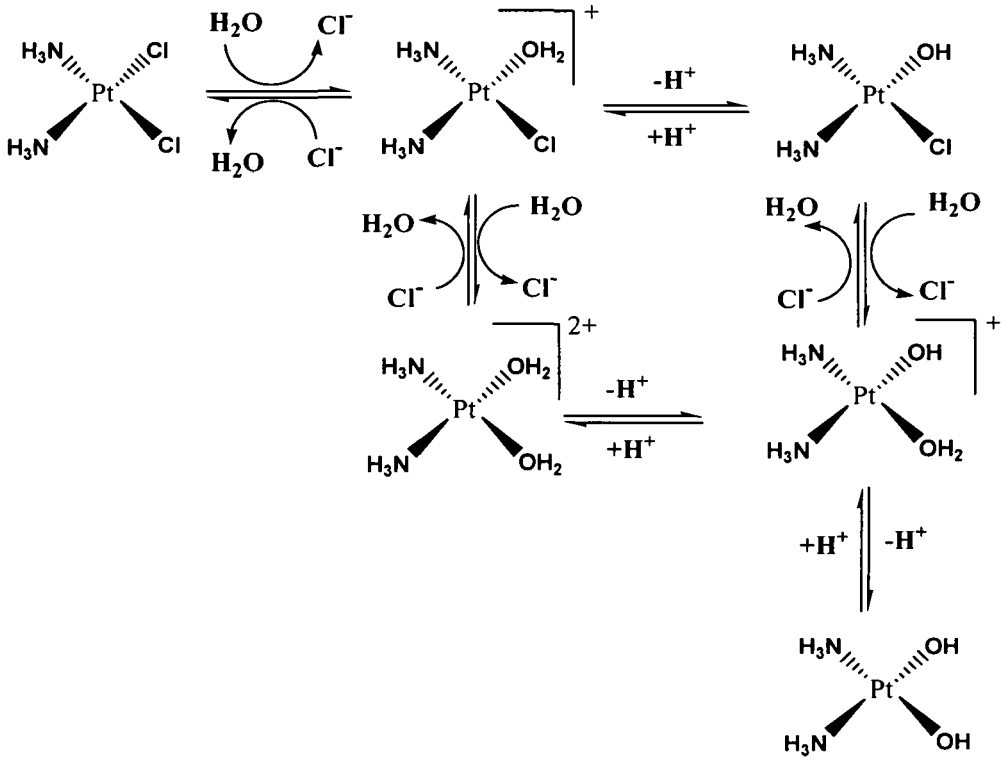


Figure 1.7: Reactions of cisplatin under physiological conditions.

These aquo species are more reactive than their parent compound towards nucleophilic centers of bio-molecules because water ligand is a much better leaving

group than chloride ligand. Inside the cell, there are many potential targets available for reactive aquo species. This includes RNA, proteins, DNA, sulfur-containing biomolecules, such as glutathione (GSH), and membrane phospholipids. The ultimate target of cisplatin is DNA, which could be identified by the filamentous growth observed in the early experiment by Rosenberg *et al.*[12] It is generally accepted that the interaction between cisplatin and DNA is largely responsible for its antitumor activity.

1.4.1.2 Cisplatin-DNA Adducts

The potential sites of DNA for platinum coordination after hydrolysis are in order of preference, the N7 atom of guanine, the N7 atom of adenine, the N1 of adenine, and N3 of cytosine. Thus N7 atom of guanine in DNA is the most accessible and reactive nucleophilic site for platinum binding to DNA. Furthermore, only N7 atoms of guanine and adenine provide directly available binding sites in the major groove of B-DNA. The binding of cisplatin with DNA results in the formation of various structurally different adducts. Initially formed monofunctional DNA adducts further react to produce variety of bifunctional intrastand and interstrand cross-links (Figure 1.8), which then block replication and prevent transcription.[25] The 1,2-intrastrand cross-links involving adjacent bases are the most abundant Pt-DNA adducts with 1,2-d(GpG) adducts comprising 60-65% of the adducts formed and 1,2-d(ApG) adducts comprising another 20-25%. In addition, there are some minor adducts including 1,3-intrastrand cross-links between nonadjacent guanines (6-8%) and interstrand cross-links (1.5%). The fact that which adducts are primarily responsible for the cytotoxicity is not yet known. However, the two most abundant adducts, i.e. intrastrand 1,2-d(GpG) and d(ApG) cross-links are believed to be responsible for the antitumor properties of cisplatin as these are not formed by the clinically inactive geometric isomer transplatin. In fact, transplatin mainly forms 1,3-intrastrand and interstrand cross-links and does not show any cytotoxicity. This hypothesis is supported by the findings that nucleotide excision repair (NER), an important DNA repair system is less effective on 1,2-intrastrand cross-links compared to 1,3-intrastrand cross-links.

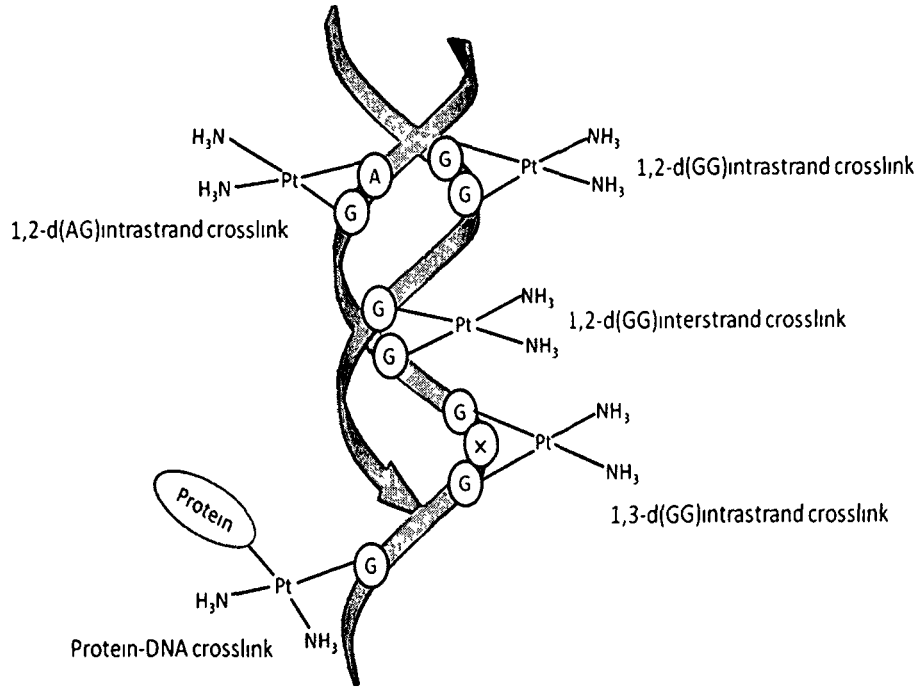


Figure 1.8: Schematic representation of cisplatin-DNA adducts.

The formation of cisplatin adducts significantly distorts the structure of target DNA resulting in a loss of helix stability and a structural change.[26,27] Intrastrand cross-links unwind the duplex DNA by bending it towards the major groove and generate a widen minor groove. Both the 1,2-d(GpG) and 1,2-d(ApG) intrastrand cross-links unwind DNA by 13°, while the 1,3-d(GpXpG) intrastrand cross-links unwind DNA by 23°. However, bending of DNA double helix is similar (32°-35°) for these three types of intrastrand adducts. Cisplatin induces a kink on DNA about 40-70° due to 1,2-d(GpG) cross-linking. The three dimensional structure of cisplatin bound to major groove of DNA through 1,2-intrastrand cross-links is shown in the Figure 1.9 (a).[19]

Several proteins are known to recognise the shallow and widen minor groove of the cisplatin-DNA adduct, such as DNA repair proteins, histones, and high mobility group (HMG) domain proteins. The HMG domain proteins specifically recognize the 1,2-intrastrand adducts and enhance the DNA bending (Figure 1.9 (b)). Interestingly, these proteins are unable to recognized the 1,3-d(GG) intrastrand adducts which further support the 1,2-intrastrand cross-links as the main adducts responsible for the antitumor activity of cisplatin.

a)

b)

Figure1.9: a) Structure of the 1,2-d(GG) intrastrand DNA-cisplatin adduct. b) Structure of domain A of HMG1 protein bound to a cisplatin 1,2-d(GG) intrastrand adduct.

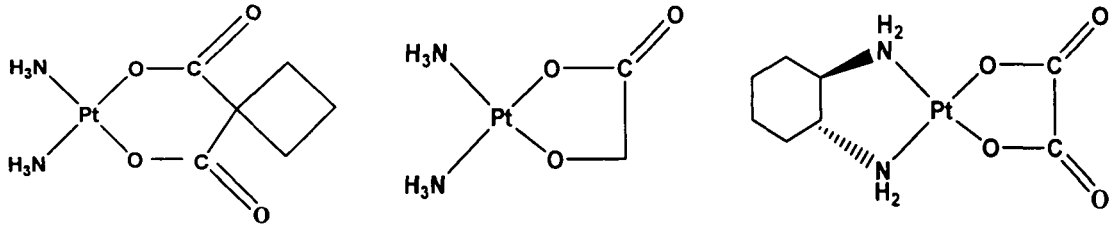
1.4.2 Disadvantages of Cisplatin

Although cisplatin has shown to be very effective in treating different kinds of cancer cell lines with cure rate close to 95% against testicular cancer, the drug does have some disadvantages. The critical drawbacks of cisplatin include its intrinsic and acquired cellular resistance, limited solubility in aqueous solution, and other toxic side effects such as nausea, ear damage, vomiting, and nephrotoxicity,

the latter can be reduced through hydration therapy and diuresis.[28] A number of new potential drug candidates were designed in order to overcome toxicity and resistance.

1.4.3 Development of New Platinum Anticancer Drugs

Over the years, various platinum complexes have been synthesized in an attempt to overcome the problems of cisplatin. Majority of these compounds were synthesized under the set of 'Structure-Activity Relationship' (SAR) summarized by Cleare and Hoeschele.[29,30] This relationship states that platinum complexes must possess two leaving groups in *cis* geometry to show anticancer activity with the general formula *cis*-[Pt X_2 $(Am)_2$], where X is the anionic leaving ligand and Am is inert amine carrier ligand. The opposing Am ligands with at least one N-H moiety are found to play important role. A carrier ligand targets a platinum drug to the cancer cell, facilitates the uptake and increases the affinity for DNA. Several platinum complexes which obey these SAR have entered the clinical trials. The second generation platinum drug *cis*-diammine-1,1-cyclobutane-dicarboxylato-platinum(II) (carboplatin; 2, Figure 1.10) has received worldwide approval in a routine clinical use.[31] The observed pharmacokinetic differences between cisplatin and carboplatin depend on the slower rate of conversion of carboplatin to the reactive species. Replacement of the chloride groups in cisplatin by cyclobutane-dicarboxylate ligand significantly diminished the nephrotoxic effects of the formed carboplatin without affecting its antitumor activity.

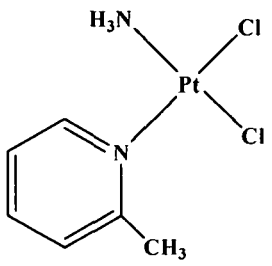


2 3 4

Figure 1.10: Molecular structure of cisplatin analogues used in the clinic: carboplatin (**2**), nedaplatin (**3**), and oxaliplatin (**4**)

Other drugs containing oxygen ligands as leaving groups include *cis*-diammine(glycolato)platinum(II) (nedaplatin; 3, Figure 1.10)[32] and 1R,2R-diamino-cyclohexane)oxalatoplatinum(II) (oxaliplatin; 4, Figure 1.10).[33] Nedaplatin was approved for clinical use in Japan since 1994.[34] It exhibits good anticancer activity and reduced toxicity.[35,36] Oxaliplatin was approved in Germany, France, and USA mainly for the secondary treatment of metastatic colorectal cancer.[37] The third generation drug oxaliplatin contains 1,2 diaminocyclohexane as carrier ligand, which increases the lipophilicity of the drug as well as improves its uptake and makes the drug more effective in inhibiting DNA chain elongation. Unfortunately these second and third generation drugs also suffer from drug resistance and other side effects.[14,38]

In search of new platinum drugs which are less toxic and can show activity against cisplatin resistance cell lines, a strategy was tried to design new complexes with decreased reactivity. It includes Pt(II) complexes with sterically hindered bulky planar ligands, such as pyridine and substituted pyridine. They are expected to be less susceptible to deactivation by sulfhydryl groups prior to DNA binding without affecting the cytotoxic activity.[39,40] One of the most prominent complex in this group is *cis*-amminedichloro(2-methylpyridine)platinum(II) (AMD473 or ZD0473; 5, Figure 1.11). AMD473 exhibited no cross-resistance to cisplatin in *in vitro* tests against human ovarian carcinoma cells while producing marked activity in both murine and human ovarian carcinoma cell lines.[41] In 1997, this orally active drug entered clinical trials and showed response in a variety of solid tumors, including ovarian, lung, and hormone-refractory prostate cancer.[42,43] The drug has an acceptable safety profile with less toxicity to the kidney and peripheral nervous system than certain other marketed platinum drugs.

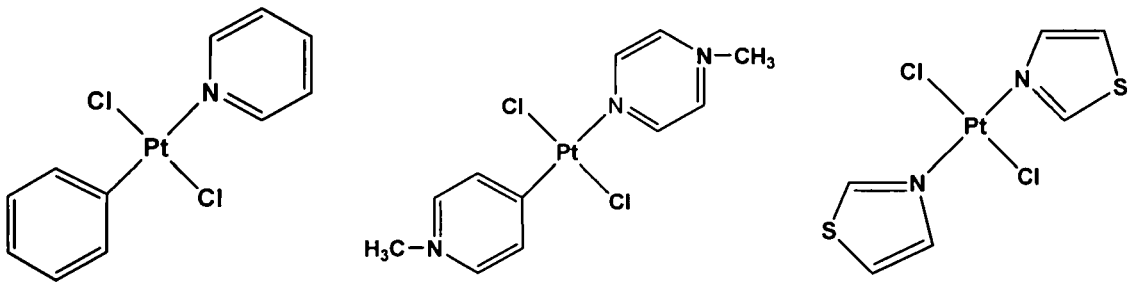


5

Figure 1.11: Molecular structure of sterically hindered platinum(II) complex AMD473.

Over the past decades several non-classical platinum complexes were found which do not obey the postulated SAR. These include active *trans* platinum (II) complexes, platinum (IV) complexes, and polynuclear platinum complexes. These classes of compounds do, however, display high cytotoxicity, both in cisplatin sensitive and cisplatin resistant cell lines.

Although transplatin, the *trans* isomer of cisplatin is not antitumor active, the exception was first reported in 1989 showing that *cis* orientation of the ligands is not a prerequisite for antitumor activity.[44] The *trans* isomers form high proportion of interstrand cross-links upon interaction with DNA, which is believed to be the reason behind their success over cisplatin resistance in certain tumors.[45,46] Figure 1.12 shows some active *trans* platinum complexes with general formula, $[PtCl_2 (L) (L')]$, where L and L' are pyridine (6), *N*-methylimidazole (7), and thiazole (8).

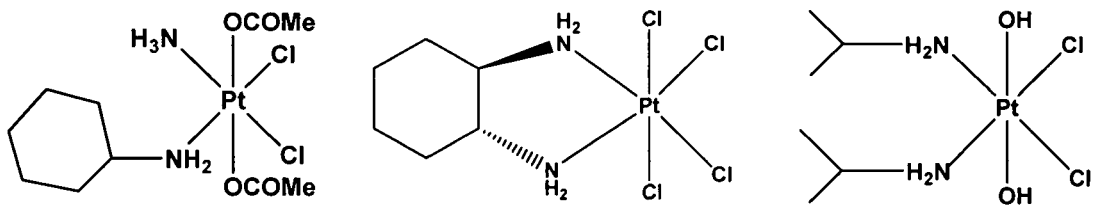


6 **7** **8**

Figure 1.12: Molecular structure of three known active *trans* platinum complexes.

It has been found that platinum(IV) alkylamines with axial carboxylate groups show selective cytotoxicity to cisplatin-resistant human tumor cell lines.[47] The most successful drug of this category is *bis*-(acetato) amminedichloro-(cyclohexylamine)platinum(IV) (satraplatin, JM216; 9, Figure 1.13) which is first orally administrative drug. This complex shows no cross-resistance with cisplatin and carboplatin and has relatively mild toxicity profile.[48] Figure 1.13 presents the two active Pt(IV) drugs, ormaplatin, (10) and iproplatin (11). The biological response of these complexes is due mainly to the lipophilicity of the axial groups combined with activation of the compalex *via* reduction to the platinum(II) species.

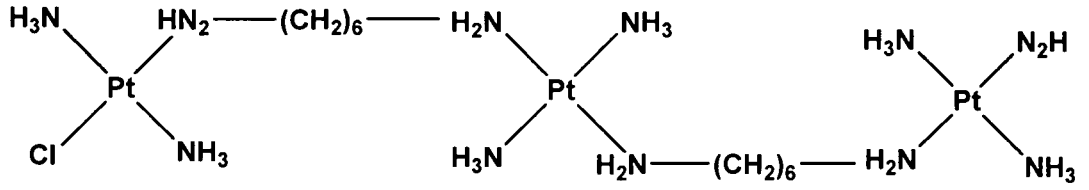
The axial lipophilic groups facilitate intestinal absorption of the drug, making oral administration possible.



9 **10** **11**

Figure 1.13: Molecular structure of the anticancer platinum complexes satraplatin or JM216 (**9**), ormaplatin, (**10**), and iproplatin (**11**).

In continuous search of new platinum compounds that form different types of cross-links with DNA, several dinuclear (and polynuclear) complexes have been studied.[49] These complexes produce long-range inter and intrastrand cross-links which are not recognized by DNA repair proteins.[50,51] Each of the two or more linked platinum centres of polynuclear complexes bind with DNA to form different structures of DNA, that are drastically different from the distortions induced by cisplatin. Farrell and co-workers have extensively investigated various di- and trinuclear platinum complexes and reported their antitumor activity *in vitro* and *in vivo*.[52] Many of these complexes are characterized by their lack of cross-resistance with cisplatin and has low nephro and neurotoxicity.[53] Recently, several binuclear,[54] trinuclear[55] and tetranuclear[56] platinum complexes have been reported. The most important of these complexes is the trinuclear Pt compound BBR 3464 (12) which is shown in Figure 1.14.



12

Figure 1.14: Molecular structure of trinuclear platinum(II) complex BBR-3464.

Thus the search for new platinum-based drugs, with an important goal of overcoming inherent and acquired resistance to the treatment with cisplatin has led to the synthesis and testing of many new platinum complexes. Discovering a new drug is a complex, expensive, and very time consuming task, as there is no single systematic way to automatically discover a drug even though the disease, targets, and molecular mechanism of drug activity are well understood.[57] The rational drug design strategies, especially the *in silico* (virtual screening)-based quantitative structure-activity relationship (QSAR) modelling techniques have emerged as a promising alternative or complementary tool toward the effective screening of potential drugs.

1.5 Nucleoside Analogues

Nucleosides are fundamental building blocks of biological systems which are thought to interact with various intracellular targets involved in the metabolism of DNA synthesis. Nucleoside analogues, which are good substrates for cellular kinases, but resistant to other host enzymes such as phosphorylases, are essential for the development of useful therapeutic agents. A large number of nucleoside analogues have been synthesised and their cytotoxic activity against a panel of tumors, in particular against murine leukemia L1210/0 cell line has been studied in the past.[58] However, nucleoside anticancer drugs are commonly associated with various adverse effects. Apparently, there is a need to search for nucleoside analogues that can selectively inhibit cancer cell proliferation. For instance, it has been seen that 5-substituted-2'-deoxyuridines reduce cancer cell proliferation by inhibiting thymidylate synthase, an eznzyme essential in the synthesis of DNA. With the aim of overcoming the drawbacks, several analogues have been synthesized in which the furanose ring has been replaced by a different five-membered ring system including carbocycles. Pioneering examples of this new generation of compounds are carbovir,[59] lobucavir,[60] dioxolane T,[61] or lamivudine[62] (Figure 1.15). The replacement of the oxygen in the sugar portion of the nucleoside with a methylene (CH_2) unit results in carbocyclic nucleoside analogues which makes the molecules more resistant to hydrolases than the natural nucleosides. Carbocyclic analogues of nucleosides have attracted growing interest due to their stability *in vivo* and powerful antitumor activities of some of these compounds.

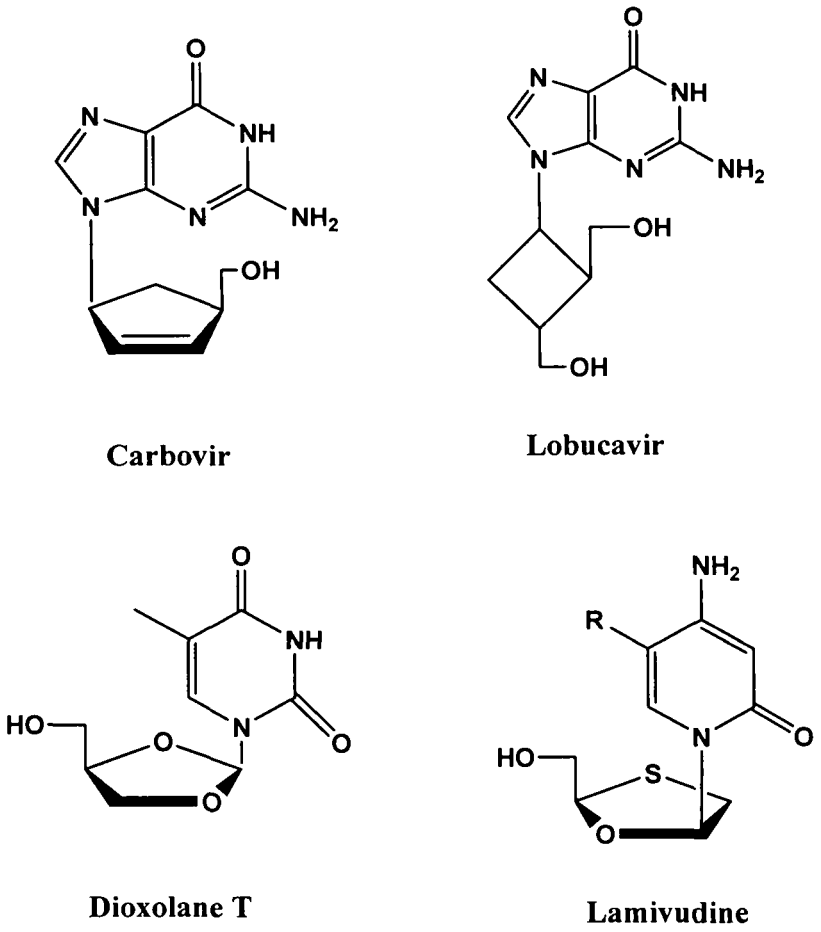


Figure 1.15: Few examples of nucleoside analogues.

The anticancer mechanism of nucleoside analogues is still not well understood at the molecular level. This is a very complex problem that certainly benefits from a good understanding of the chemical properties at the molecular level such as steric, lipophilic, and electronic characteristics by using more versatile computational techniques, like QSAR.

1.6 Computer-Aided Drug Design

The design of drugs is an extremely complex process. Although experimental screenings can be used to explore active chemical structure, it cannot be carried out on millions of candidate molecules due to prohibitive costs both in terms of time and money. In the last decade new tools have become available for drug design including computational chemistry, high-throughput screening and combinatorial chemistry. With *in silico* approaches, it is possible to reduce the amount of

compounds for experimental evaluations. The *in silico* approaches that include QSAR techniques are increasingly attracting the attention of medicinal chemists as well as the pharmaceutical industry.[63-67] In Figure 1.16 we have shown the increasing number of publications on computational drug design in 5 year increments.

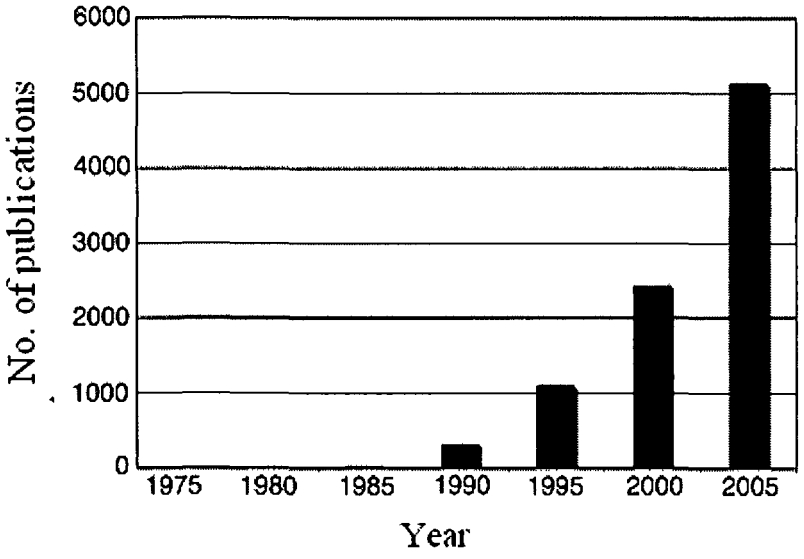


Figure 1.16: Number of publications on computational drug design in 5 year increments (taken from ISI Web of Science).

Computer aided drug design has matured into a scientifically still challenging, but industrially highly applicable field that has achieved considerable successes.[68-70] It employs the combination of knowledge of molecular interactions and drug activity, together with detailed energy calculations, and geometric considerations in an effort to narrow down the search for effective drugs. The computational chemistry tools include empirical molecular mechanics, semiemperical methods, Monte Carlo, molecular dynamics, and *ab-initio* quantum chemical methods. Effect of solvent can also be incorporated along with many of these various methods. The foundation of computational approaches to drug design is the correlations of computed and experimentally evaluated properties of molecules. If the computed properties are based on a more rigorous theory such as modern quantum chemistry, their accuracy competes with or even surpasses the accuracy of experimental measurements.[71] However, the role of quantum mechanical methods has been until now very limited in the drug design process, mostly

because of the high computational demands involved that allowed to deal only with small molecules.[72] Nevertheless, the recent progress in first-principles electronic structure calculations together with the steady increase in computational power has considerably broadened the range and scope of application of these theoretical methods. Of particular interest are density functional theory (DFT) calculations that have been proven to be a powerful tool for studying a large variety of problems in chemistry and more recently in highly complex systems of biophysics and biochemistry.

In the present thesis, density functional theory is used to perform QSAR analysis of several drug molecules. The drug-DNA interaction has been studied using Hartree-Fock and DFT along with quantum mechanics/molecular mechanics (QM/MM) technique. All these computational methods are briefly discussed in the following section.

1.7 Molecular Mechanics

Molecular mechanics is based on simple classical model of molecular structure where atoms are treated as hard spheres. The principle of molecular mechanics is to express the energy of a molecule as a function of its resistance toward bond stretching, bond bending, and atom crowding, and to use this energy function to get various possible potential energy surface minima. The interactions between the atoms in a molecule are described by force field method.

The potential energy expression of a molecule in the force field can be written as

$$E = \sum_{bonds} E_{stretch} + \sum_{angles} E_{bend} + \sum_{dihedrals} E_{torsion} + \sum_{pairs} E_{nonbond} \qquad (1.1)$$

in which :

$E_{stretch}$ is the energy of a bond stretched or compressed from its natural bond length.

E_{bend} is the energy of bending bond angles from their natural values.

$E_{torsional}$ is the torsional energy due to twisting around bonds.

$E_{nonbond}$ is the energy due to nonbonded interactions.

The approximations adopted in molecular mechanical calculation make the computations inexpensive. The speed of molecular mechanics enables large biomolecular systems containing thousands of atoms, such as proteins and DNA to be treated. For system where good parameters are available, it is possible to predict accurate geometries and relative energies for a larger number of molecules in a short time. The weaknesses of this method stem from the facts that it ignores electrons and lacks good parameters. This method is applicable only for a limited class of molecules for which the force field is parameterized. Also, molecular mechanics cannot provide information about the shapes and energies of molecular orbitals and cannot throw light on phenomena such as electronic spectra.

1.8 Quantum Mechanics

The electronic structure and total energy of atoms, molecules and crystals can be obtained by solving Schrödinger eigen value equation. There are several formulations of this equation, but the one concerned in the thesis is the time-independent, non-relativistic Schrödinger equation:

$$\hat{H}\,\Psi(\mathrm{r}) = E\,\Psi(\mathrm{r}) \qquad (1.2)$$

where $\hat{H}$ is the Hamiltonian operator for a molecular system and can be written as the kinetic and potential energies of the nuclei and electrons:

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \qquad (1.3)$$

where T is the kinetic energy operator, V_{ne} is the external potential produced by the nuclei acting on the electrons, and the rest two terms, V_{ee} and V_{nn} are the electron-electron and nuclei-nuclei potentials, respectively.

The solution to the Schrödinger equation thus involves finding the correct wave function of the system. Since an infinite number of wave functions can be constructed which satisfies eq. (1.2), one should proceeds systematically to get the wave function of the ground state Ψ_0, i.e the state which delivers the lowest energy E_0. In this context, the term variational principle comes which holds a very prominent place in all quantum-chemical applications. It states that, for a ground state wave function, any trial wave function will have an energy eigen value higher or equal to the exact energy of the ground state (corresponding to the exact wave function), i.e.,

$$\left\langle \psi_{trial} \middle| \hat{H} \middle| \psi_{trial} \right\rangle = E_{trial} \geq E_0 = \left\langle \psi_0 \middle| \hat{H} \middle| \psi_0 \right\rangle \tag{1.4}$$

ψ_{trial} is the trial wave function which is usually a linear combination of basis functions:

$$\psi_{trial} = \sum_n c_n \psi_n \tag{1.5}$$

where ψ_n are orthonormal wave functions, such that $\sum |c_n|^2 = 1$

For atoms, the Schrödinger equation can only be solved analytically for systems with one electron. The solution of the Schrödinger equation becomes increasingly more complex when more electrons are present. Consequently approximation must be made. One of the most important approximation is the Born-Oppenheimer approximation. In this approximation the motion of electrons and nuclei are separated by expanding the total molecular wave function as a product of electronic and nuclear wave functions. Since the electrons are by orders of magnitude lighter than the nuclei, they move much faster and relax very rapidly to the instantaneous ground-state configuration given by the nuclear positions. This allows us to calculate the wave function for electrons moving in the potential field of the nuclei, treated as fixed point charges:

$$\Psi(r_i, R_\alpha) = \Psi(r_i, R_\alpha)\ \phi(R_\alpha) \tag{1.6}$$

Thus the complete Hamiltonian given in equation (1.3) reduces to the so-called electronic Hamiltonian

$$\hat{H}_{elec} = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2\right) + \sum_{i=1}^{N} v(r_i) + \sum_{i<j}^{N} \frac{1}{r_{ij}} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} \tag{1.7}$$

in which

$$v(r_i) = -\sum_\alpha \frac{Z_\alpha}{r_{i\alpha}} \tag{1.8}$$

The electronic wave function $\Psi(r_i, R_\alpha)$ depends on the electron coordinates, while the nuclear coordinates enter only parametrically.

Quantum chemical methods are divided into two classes: semi-empirical methods (such as CNDO, MNDO etc.) and non-empirical (*ab initio*, DFT etc.) methods. Semi-empirical methods use some parameters derived from experimental data to simplify the calculations and it is less demanding than *ab initio* methods. An

a*b initio* calculation uses the correct Hamiltonian and does not use the experimental data other than the values of some physical constants such as speed of light, the masses, charges of the electrons and nuclei, Plank's constant etc.

1.8.1 The Hartree-Fock Method

The Hartree-Fock (HF) method is the most common *ab-initio* method. It approximates the N-electron wave function as an antisymmetrized product (the Slater determinant, Φ_{SD}) of N one-electron wave functions $\psi_i(x_i)$. Each electron moves independently in the spin orbital space and it experiences a Coulombic repulsion due to the average positions of electrons. It experiences exchange interaction due to antisymmetrization.

$$\psi_{HF} \approx \Phi_{SD} = \frac{1}{\sqrt{N!}} \det[\psi_1 \psi_2 .... \psi_N] \tag{1.9}$$

In the Hartree-Fock approach the spin orbitals ψ_i are varied under the constraint that they remain orthonormal such that the energy obtained from the corresponding Slater determinant is minimal:

$$E_{HF} = \min_{\Phi_{SD} \to N} E[\Phi_{SD}] \tag{1.10}$$

The derivation of the expectation value of the Hamiltonian operator with a Slater determinant finally gives the HF energy expression:

$$E_{HF} = \left\langle \psi_{HF} \middle| \hat{H} \middle| \psi_{HF} \right\rangle = \sum_{i=1}^{N} H_i + \frac{1}{2} \sum_{i,j=1}^{N} (J_{ij} - K_{ij}) \tag{1.11}$$

where

$$H_i = \int \psi_i^*(x)[-\frac{1}{2}\nabla^2 + v(x)]\psi_i(x)dx \tag{1.12}$$

defines the contribution due to the kinetic energy and electron-nucleus attraction and

$$J_{ij} = \int\int \psi_i(x_1)\psi_i^*(x_1)\frac{1}{r_{12}}\psi_j^*(x_2)\psi_j(x_2)dx_1 dx_2 \tag{1.13}$$

$$K_{ij} = \int\int \psi_i^*(x_1)\psi_j(x_1)\frac{1}{r_{12}}\psi_i(x_2)\psi_j^*(x_2)dx_1 dx_2 \tag{1.14}$$

are Coulomb and exchange integrals, respectively.

For a polyatomic system the most common approach for the variational analysis is to determine a linear combination of atomic orbitals (called Molecular Orbitals, MOs or sometimes MO-LCAO) that minimize the total electronic energy

under the constraint that the MOs remain orthonormal. Minimization of eq. (1.11) with the constraint that $\psi_i(x_i)$ remain orthonormal, gives the Hartree-Fock equation which introduces the Lagrangian multipliers ε_i that have the physical interpretation of orbital energies:

$$\hat{F}\psi_i(x) = \varepsilon_i\psi_i(x) \tag{1.15}$$

The fock operator $\hat{F}$ is an effective one-electron operator defined as

$$\hat{F} = -\frac{1}{2}\nabla^2 + v + \hat{g} \tag{1.16}$$

in which the Coulomb-exchange operator $\hat{g}(x_1)$ has the following two components:

$$\hat{g} = \hat{j} - \hat{k} \tag{1.17}$$

Here, the Coulomb operator $\hat{j}$ and exchange operator $\hat{k}$ are defined as

$$\hat{j}(x_1)f(x_1) = \sum_{k=1}^{N}\int \psi_k^*(x_2)\psi_k(x_2)\frac{1}{r_{12}}f(x_1)dx_2 \tag{1.18}$$

and

$$\hat{k}(x_1)f(x_1) = \sum_{k=1}^{N}\int \psi_k^*(x_2)f(x_2)\frac{1}{r_{12}}\psi_k(x_1)dx_2 \tag{1.19}$$

with $f(x_1)$ an arbitrary function.

The Hartree-Fock equations form a set of pseudo-eigen value equations. A specific Fock orbital can only be determined if all the other occupied orbitals are known, and iterative methods must therefore be employed for determining the orbitals. The method used is called self-consistent field (SCF) method which starts with a guess set of orbitals. The resulting new set of orbitals is then used in the next iteration and so on until the input and output orbitals differ by less than a predetermined threshold.

1.8.2 Basis Sets

A basis set is a mathematical description of orbitals of a system, which is used to approximate theoretical calculation or modelling. Two types of basis sets dominate the area of *ab initio* calculations. Those are Slater Type Orbitals (STO) and Gaussian Type Orbitals (GTO). Slater type orbitals are simple exponentials that mimic the exact eigenfucntions of the hydrogen atom and have the functional form:

$$f^{STO}(r) = \left(\frac{\xi^3}{\pi}\right)^{0.5} exp(-\xi r) \qquad (1.20)$$

Slater functions are good approximations to atomic wave function. However, for Hartree-Fock SCF calculations on molecules with three or more atoms, the evaluation of the many electron integrals is important. The introduction of GTOs to replace STOs in calculations made the evaluation of three and four centre integrals more rapid. All most all *ab initio* calculations nowadays use Gaussian orbitals of the form

$$f^{GTO}(r) = \left(\frac{2\alpha}{\pi}\right)^{0.5} exp(-\alpha r^2) \qquad (1.21)$$

The main advantage of GTOs is that the product of two Gaussians at different centers is equivalent to a single Gaussian function centered at a point between the two centers. Thus two electron integrals on three and four different atomic centers can be reduced to integrals over two different centers, which are much easier to compute.

The simplest type of basis set is minimal basis set which includes one basis function for each SCF occupied atomic orbital with distinct principal and angular quantum numbers. A significant improvement is made by adopting a double-zeta (DZ) and triple-zeta (TZ) basis sets in which each basis function in the minimal basis set is replaced by two and three basis functions, respectively. A split valence basis set is a compromise between the inadequacy of a minimal basis set and the computational demand of DZ and TZ basis sets which use DZ-type basis set for the valence atomic orbitals and a minimal description for the core orbitals. During bond formation, atomic orbitals are distorted or polarized from their idealized atomic shapes. Orbital polarization may be mimicked by including polarized basis functions that incorporates the basis functions of higher angular quantum number. In calculations involving electronically excited state or having electron density away from the nuclear centers, diffuse functions are often added to the basis set.

1.8.3 Effective Core Potentials (ECP)

These basis functions are useful for heavy elements (4^{th} period onwards) to include *relativistic effects*. The electrons near the very positive nucleus of a heavy element experience a larger relative attraction than for lighter elements, which

accelerates the electrons close to the speed of light. In this situation, Einstein's theory of general relativity starts to have an effect on the shape of the atomic orbitals. The net effect is that the inner core orbitals of a heavy element are contracted relative to the corresponding orbitals in a lighter element. A specific type of basis set that has been used in this work is the Effective Core Potential (ECP). The ECP basis set includes valence electrons explicitly and replaces the influence of the core electrons by a static, effective potential, commonly described by a polynomial function. The effective potential accounts for relativistic effects, and tremendously reduces the computational expense for heavier elements where relativistic effects are significant for the core electrons of the atoms.

1.8.4 Electron Correlation

Since electrons repel, they will try to avoid each other. Thus the motion of each electron is correlated with the motion of the other electrons in the system. Within the HF formalism the antisymmetric wavefunction is approximated by a single Slater determinant, which does not include this Coulomb correlation. Therefore the energy calculated with HF theory is different from the exact energy of the system. The difference between these two is called the correlation energy

$$E_{corr} = E_{exact} - E_{HF} \qquad (1.22)$$

The development of methods to determine the correlation contributions accurately and efficiently leads to a broad categories of approaches called post HF methods. These include the linear mixing of many determinants called configuration interaction (CI), coupled cluster (CC), and Møller-Plesset perturbation theory (MP2, MP3, MP4, etc.). Usually, post-Hartree-Fock methods give more accurate results than Hartree-Fock calculations, although the added accuracy comes with the price of added computational cost.

1.9 Density Functional Theory (DFT)

Density functional theory provides an expression for the ground state energy of a system of interacting electrons in an external potential as a functional of the ground state electronic density. This theory is different from the traditional quantum chemical methods based on the wavefunctions. The fundamental underlying mathematics of this method is the *functional* which is defined as a function of a

function. The electronic energy of a system may be written as a function of the electron density $\rho(r)$ and thus it is said to be functional of ρ, E[ρ]. This means that for a given density function $\rho(r)$ there is one and only one energy value. The electron density is defined as

$$\rho(r) = N\int....\int\lfloor\Psi(x_1x_2....x_n\rfloor^2 dx_1dx_2...dx_n \qquad (1.23)$$

$\rho(r)$ determines N, the total number of electrons, *via* its normalization $\int \rho(r)dr = N$. Clearly ρ(r) is a non-negative function of only three spatial variables in contrast to wave function (ψ), which is a function of $4N$ variables (3 spatial variables and 1 spin variable for each electron) for an N electron system. Moreover, unlike the wave-function, the electron density is an observable and can be measured experimentally, e.g. by X-ray diffraction.[73]

Density Functional Theory is based on two theorems proposed by Hohenberg and Kohn (HK) in the early 60s.[74] The first theorem proved that there can be no more than one external potential $v(r)$ for a given electron density. Since ρ determines N and v, hence the ground state energy and all the properties of the ground state of a system are uniquely defined by the electron density. In other words, the ground state energy is a unique functional of density ρ with the following components

$$E_0[\rho] = T_e[\rho] + V_{ee}[\rho] + V_{ext}[\rho] \qquad (1.24)$$

V_{ext}, the energy corresponding to the external potential, depends on the system while the kinetic energy of electorns, T_e and energy of interaction among electrons, V_{ee} are independent of the system.

The Eq. (1.24) is modified as

$$E_0[\rho] = \int \rho(r)v(r)dr + F_{HK}[\rho] \qquad (1.25)$$

where the Hohenberg-Kohn functional $F_{HK}[\rho] = T_e[\rho] + V_{ee}[\rho]$

The second of HK theorem states that the variational principle is also valid for the electron density. The true ground state energy E_0 is always less than the energy calculated with any other density, ρ'. i.e. $E_0[\rho] \leq E_0[\rho']$, the equality holds only $\rho = \rho'$

Unfortunately, these theorems offer no guide about the exact ground state density ρ and the exact Hohenberg-Kohn functional F_{HK}. This problem was addressed by Khon and Sham in 1965.[75] They separated the energy functional of eq. (1.24) as follows

$$E[\rho] = T_S[\rho] - \sum_A \int \frac{Z_A}{R_{1A}} \rho(r_1) dr_1 + \frac{1}{2} \iint \rho(r_1)\rho(r_2) \frac{1}{r_{12}} dr_1 dr_2 + E_{XC}[\rho] \qquad (1.26)$$

where $T_S[\rho]$ is the kinetic energy of electrons in a system with non interacting electrons. Second term describes the potential effected from the nuclei. Third term is the purely Coulombic repulsion between the electrons. The last functional, $E_{XC}[\rho]$ is called the exchange-correlation energy which contains the non-classical electrostatic interaction energy and the difference between the kinetic energies of the interacting and non-interacting systems. This quantity is not solvable in this form.

Within the Kohn-Sham (KS) formalism, total density analogous to the wave function in HF formalism, consists of orbitals, i.e.

$$\rho(r) = \sum_i |\phi_i(r)|^2 \qquad (1.27)$$

This leads finally to the Kohn-Sham equations

$$f^{KS}\phi_i(r) = \varepsilon_i \phi_i(r) \qquad (1.28)$$

where ε_i is the KS orbital energy and f^{KS} is the KS operator

$$f^{KS} = \left\{ -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{R_{1A}} + \int \frac{1}{r_{12}} \rho(r_2) dr_2 + V_{XC}[r_1] \right\} \qquad (1.29)$$

This operator is similar to the Fock operator. Hence eq. (1.28) can be solved by same algorithm as in the HF theory. The Khon-Sham orbitals $\phi_i(r)$, which can be easily derived from this equation, can be used immediately to compute the total density using eq. (1.26) leading to a new cycle of self consistent field.

Unfortunately, although the theory unlike HF is in principle exact, the energy functional contains the unknown quantity, $E_{XC}[\rho]$ that must be approximated in any practical implementation of the method. If the functional form of $E_{XC}[\rho]$ and consequently the exchange correlation potential, were available, one could solve the N-electron problem by finding the self-consistent solution of a set of single-particle equations.

1.9.1 Approximate Exchange-Correlation Functionals

Several approximate exchange-correlation functionals have been proposed in the literature, the most commonly used ones being the local density approximation (LDA), the generalized gradient approximation (GGA) and, more recently, the hybrid functional. LDA functional have been derived by using the assumption that the electron density in some point *r* locally could be treated as a uniform electron gas. The local density approximation[76] is the simplest functional:

$$E_{XC}^{LDA}[\rho(r)] = \int \varepsilon_{XC}^{hom\,o}(r)[\rho(r)]\rho(r)dr \qquad (1.30)$$

Although this might not seem appropriate for real atoms and molecules, the LDA method is often surprisingly accurate and gives generally very good results for systems with slowly varying charge densities.

GGA functional[77] emerged as a correction to the LDA in which the correction term, which is a function of the density gradient in point *r*, is added to the LDA functional:

$$E_{XC}^{GGA}[\rho(r)] = \int \varepsilon_{XC}^{GGA}(r)[\rho(r)]\rho(r)dr + \int F_{XC}\left[\rho(r),|\nabla\rho(r)|\right]dr \qquad (1.31)$$

where F_{XC} is a suitable functional, which satisfies certain known limits of E_{XC}.

Many different GGA's are available in the literature. One of the most popular functionals is given by the combination of the Becke exchange functional and the Lee, Yang, and Parr correlation functional (BLYP)[78,79] In this work we mainly use this functional, which has shown to give in general very good results for biological systems.[80,81]

In recent years, hybrid functionals have become very popular in particular for chemical applications. These functional are linear combinations of the exact exchange contributions (from HF) and DFT exchange-correlation functional:

$$E_{XC}^{hybrid}[\rho(r)] = E_{XC}^{GGA}[\rho(r)] + C_X\left(E_X^{HF}[\rho(r)] - E_X^{GGA}[\rho(r)]\right) \qquad (1.32)$$

where $E_{XC}^{GGA}[\rho(r)]$ and $E_X^{GGA}[\rho(r)]$ are exchange-correlation and GGA exchange energies, and $E_X^{HF}[\rho(r)]$ is the exact exchange which has the same form as the HF

exchange energy. The most popular hybrid functional, the three parameter B3LYP functional[82] combines LDA and the BLYP GGA with exact exchange.

1.10 DFT-Based Reactivity Descriptors

Several reactivity descriptors are used in DFT to define the reactivity of molecules. DFT provides a framework to discuss reactions in terms of changing number of electrons (N) or changing external potential, $v(r)$ due to nuclei. The first derivatives of $E(\rho)$ with respect to the number of electron N under the constant external potential $v(r)$ is termed as the chemical potential, μ. It is identified as the negative of the electronegativity (χ) by Iczkowski and Margrave[83] and defined as

$$\chi = -\mu = -\left(\frac{\partial E}{\partial N}\right)_{v(\vec{r})} \tag{1.33}$$

Global hardness (η) is defined[84] as the corresponding second derivative of energy.

$$\eta = \frac{1}{2}\left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\vec{r})} = \frac{1}{2}\left(\frac{\partial \mu}{\partial N}\right)_{v(\vec{r})} \tag{1.34}$$

Using the finite difference approximation, global hardness and chemical potential can be approximated as

$$\eta = \frac{IP - EA}{2} \tag{1.35}$$

$$\mu = -\left(\frac{IP + EA}{2}\right) \tag{1.36}$$

where IP and EA are the first vertical ionization potential and electron affinity, respectively, of the chemical system.

Further approximation using Koopmans' theorem[85], the above parameters can be expressed by taking IP and EA as negative of the HOMO and LUMO energies,

$$\mu = \frac{E_{LUMO} + E_{HOMO}}{2} \tag{1.37}$$

and

$$\eta = \frac{E_{LUMO} - E_{HOMO}}{2} \tag{1.38}$$

where E_{LUMO} is the energy of the lowest unoccupied molecular orbital and E_{HOMO} is the energy of the highest occupied molecular orbital.

Parr *et al.*[86] proposed a new reactivity descriptor, global electrophilicity index *(ω)* in terms of chemical potential and hardness as

$$\omega = \frac{\mu^2}{2\eta} \tag{1.39}$$

According to this definition, ω describes the electrophilic power of a molecule.

The site-selectivity of a chemical system cannot be studied using the global reactivity descriptors. For this purpose, appropriate local reactivity descriptors need to be defined. Local reactivity descriptors, the Fukui function (FF), a frontier MO reactivity index is by far the most important local reactivity index and defined as

$$f(r) = \left(\frac{\delta^2 E}{\delta N \delta v(r)}\right) = \left[\frac{\delta \mu}{\delta v(r)}\right]_N = \left(\frac{\delta \rho(r)}{\delta N}\right)_{v(r)} \tag{1.40}$$

Mendez and Gazquez[87] and Yang and Mortier[88] introduced a procedure to obtain information about $f(r)$. This procedure condenses the values around each atomic site into a single value that characterizes the atom in the molecule. With this approximation, the condensed Fukui function becomes

$$f_k^+ = [q_k(N+1) - q_k(N)] \quad \text{(for nucleophilic attack on the system)} \tag{1.41a}$$

$$f_k^- = [q_k(N) - q_k(N-1)] \quad \text{(for electrophilic attack on the system)} \tag{1.41b}$$

$$f_k^0 = \frac{1}{2}[q_k(N+1) - q_k(N-1)] \quad \text{(for radical attack on the system)} \tag{1.41c}$$

where $q_k(N)$, $q_k(N+1)$, and $q_k(N-1)$ are the electronic populations of the *k*th atom for *N*, *N + 1*, and *N - 1* electron systems, respectively.

Recently, Chattaraj *et al.*[89] have defined a generalized concept of philicity associated with a site *k* in a molecule as

$$\omega_k^\alpha = \omega f_k^\alpha \tag{1.42}$$

where α = +, –, and 0 represent nucleophilic, electrophilic, and radical attacks, respectively.

1.11 QSAR/QSPR

The very first computer-aided approach in drug design developed in the early 1960s, when Corwin Hansch started the QSAR (quantitative structure–activity relationship) discipline.[90,91] Since then, many new methods have revolutionized the field of "drug research". The concept has evolved from 2D-QSAR to 3D-QSAR and lately 4D-QSAR.[92] The ultimate goal of QSAR studies is to relate the biological activity (Y) of a series of compounds with some appropriate descriptors (X). There is a similar technique labelled as QSPR for "quantitative structure–property relationship, which refers the potential relationship between chemical structure and property of molecules.

The first step in creating a QSAR model is to generate a training set of similar compounds with their experimental activities (termed as dependent variable). The next step is to compute descriptors (termed as independent variables) which can be defined as numerical values that encode certain aspects of molecular structures. The relationship between descriptors and activities is

$$\text{Molecule activity} = f(\text{molecule structure}) = f(\text{descriptors}) \qquad (1.43)$$

Given a set of descriptors, a QSAR model can be built by defining a relationship between these descriptors and the observed activities with the help of statistical techniques. Figure 1.17 summarizes the process of QSAR/QSPR modelling, showing how the descriptors establish the relationship between molecular structures and biological activities or properties.

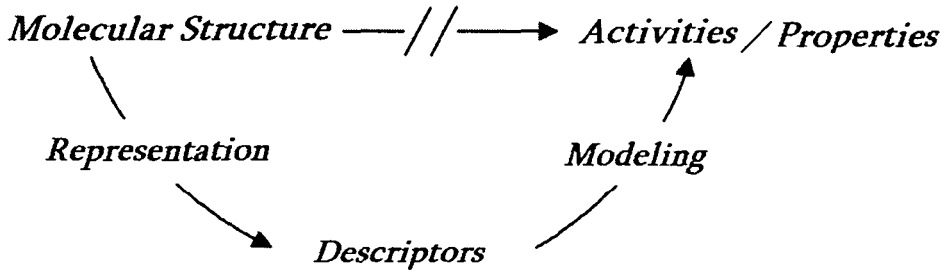


Figure 1.17: A flowchart showing the steps involved in predicting molecular properties or activities from molecular structure.

Regression analysis is used to derive a QSAR model. The conclusions drawn from a regression analysis are dependent on the assumption of the regression

model.[93] If it is assumed that the relationship is well represented by a model that is linear in the regressor variables, a suitable model may be

$$Y' = b_1 X_1 + b_2 X_2 + \ldots\ldots + b_p X_p + e \quad (1.44)$$

where $b_1, b_2 \ldots\ldots b_p$ are regression coefficients

e is the intercept

$X_1, X_2 \ldots\ldots X_p$ are independent variables

Y' represents expected values of the dependent variables by the regression model.

The objective of regression analysis is to estimate regression coefficients and the most popular algorithm to estimate the parameters is the least squares method.

To assess the accuracy of regression model squared correlation coefficient r^2 is used:

$$r^2 = \text{ESS/TSS} = 1 - (\text{RSS/TSS}) \quad (1.45)$$

where:

TSS is the total sum of squares: TSS = $\sum (Y_i - Y_{mean})^2$, it has *n-1* degrees of freedom

ESS is the explained sum of squares: ESS = $\sum Y' - Y_{mean})^2$, it has *p* degrees of freedom

RSS is residual sum of squares: RSS = $\sum (Y_i - Y'_i)^2$, it has *n-p-1* degrees of freedom.

Y_i are the observed values of dependent variable

Y_{mean} is the mean value of dependent variable

Y' are the prediced values of the dependent variables by the regression model.

n is the number of observations

p is the number of independent variables included in the regression model.

It is well known that r^2 is not always a good indicator of model quality and in many cases can be misleading. An alternative to r^2 is r^2_{CV} which is obtained by using a leave one-out (LOO) cross-validation procedure. That is, the linear model is generated using the whole dataset excluding one point.

The statistical significance of a regression equation can be assessed by means of the Fisher (*F*) statistic:

$$F = \text{EMS/RMS} \qquad (1.46)$$

where, EMS is explained mean square: EMS = ESS/*p*

RMS is the residual mean square: RMS = RSS/ *(n-p-1)*

Although the first QSAR model developed by Hansch is a linear relationship, developments in the field of statistics have produced many new methods of building predictive models, such as non linear regression techniques and algorithmic techniques. However, linear models are widely used owing to their simplicity and ease of development.

The various descriptors in use can be broadly categorized as being constitutional, topological, electrostatic, geometrical, or quantum chemical. Constitutional descriptors give a simple description of what is in the molecule. For example, the total number of atoms in the molecule, the number of double bonds, the number of rings, etc. Constitutional descriptors often appear in a QSAR equation when the property being predicted varies with the size of the molecule. Topological descriptors are numbers that give information about the atom-atom connectivity in a molecule. Some examples are Wiener index, Balaban's *J* index, Randi's molecular connectivity index, etc. Electrostatic descriptors are the ones that give information about the molecular charge distribution. Some examples are polar surface area, Mulliken atomic partial charges, dipole moment, molecular polarizability, etc. Geometrical descriptors describe the molecule's size and shape. These include molecular surface area, solvent-accessible surface area, molecular volume, moments of inertia, etc. These descriptors are often important to the QSAR equation if the property being predicted is dependent upon whether the molecule is more globular or linear. Quantum chemical descriptors consider various features of the molecules' electronic environment. Examples include HOMO and LUMO energies, electro-negativity, total energy, ionization potential, electron affinity, and various atom-centered partial charge descriptors. More recently, there is an increasing interest on density functional theory based descriptors in QSAR modelling. The importance of electrophilicity index in understanding the biological activity, especially toxicity, carcinogenicity, and mutagenicity of various biological systems has been demonstrated and found to be suitable in describing these properties effectively.[94-96]

1.12 Quantum Chemical Solvation Models

While studying biochemical processes, effects of the solvent media can not be disregarded, since chemical properties of such processes are often very different from the gas phase. There is a need for an accurate model which takes these effects into account and provides a quantitative treatment of the solvent influencing the system. Effective solvent models used in quantum chemistry may generally be divided into two main categories, the discrete solvation models and the dielectric continuum models.

In discrete model, a solute is surrounded by a number of explicitly treated solvent molecules i.e. in this method an explicit microscopic description of the solvent molecules is maintained. In order to be effective, the solute is treated with QM, whereas the solvent, or at least the major part of this, may be treated with MM. The two regions are then coupled by a QM/MM interaction Hamiltonian. This leads to the definition of the combined quantum mechanics/ molecular mechanics (QM/MM) approach.[97-99] Discrete model plays an important role in the phenomena where solvent plays an integral part in the reaction. However, this method does not take into account the long-range electrostatic interactions and polarization effects.

In continuum model, solvent media is treated as a macroscopic dielectric continuum with the sole property of providing polarization charges that mimic the average behaviour of the particular solvent of interest. In this model the solvent assumes linear response to a perturbing electric field. The model was developed and exploited by Born,[100] Onsager[101] and Kirkwood.[102] In cases where specific solute-solvent interactions are of no consequence to the chemistry of the system, continuum model provides a computationally efficient means to model the influence of the solvent. In addition, continuum model automatically gives a configurationally sampled solvent effect, and this allows one to avoid statistical analyses based on delicate sampling averages.

The most successful solvent model so far is the polarizable continuum model (PCM) by Tomasi and co-workers.[103] In the PCM model the solvent is represented by a homogeneous dielectric continuum of permittivity ε in which a cavity is created approximating the shape of the solute to accommodate the species under study. The cavity is obtained by assigning a sphere of given radius to each atom forming the solute and considering the final union of these interlocking spheres. The dielectric

continuum is polarized by the solute that causes a charge distribution on the surface of solute cavity. For a given charge distribution within the cavity, the surface screening charges are calculated by solving the Poisson equation subject to cavity boundary conditions.[104,105] The calculated screening charges are then used as a perturbation included in the Hamiltonian for electronic structure calculations of the solute.

Another solvent model which is very similar to PCM is the conductor-like screening model (COSMO) developed by Klamt and his co-workers[106] and used later by Truong *et al.*[107] This model assumes the surrounding medium as a conductor, which simplifies the electrostatics computations and allows for the calculation of accurate gradients without cavity shape constraints. The surface charges are calculated directly from the electrostatic potential of the charge distribution within the cavity. This approach leads to better numerical accuracy, and is simpler, particularly for the analytic gradient evaluation.

1.13 Quantum Mechanics/Molecular Mechanics (QM/MM) Methods

Nowadays, theoretical calculations of molecular and electronic structure represent a valuable complement to experiments to elucidate structure–activity relationships of drugs and to study their binding mechanism at the molecular level. The first principle based QM methods can be employed successfully in electronic structure calculations of small models of drug molecules. However, these methods are computationally quite expensive and their treatment is usually restricted to relatively small systems in the gas phase. Finally, the development of hybrid QM/MM schemes enable the study of complex biochemical processes occur in a heterogeneous condensed phase environment comprising several thousands of atoms with high precision. In these schemes the region of biological interest, such as active site of protein or the binding site in drug-DNA complex is treated by high level QM while remaining protein or DNA residues as well as the solvent and counter ions are treated by MM. ONIOM (our Own N-layer Integrated molecular Orbital molecular Mechanics) is a hybrid QM/MM method developed by Morokuma and co-workers[108-111] and this method has been implemented in Gaussian03 program. It is an efficient method in handling the QM/MM computations owing to its

flexibility in allowing combinations of different methods for different parts of the molecular system. The energy calculation for two-layer ONIOM method is based on three separate calculations:

$$E^{ONIOM\ 2} = E^{low}_{real\ system} + E^{high}_{model\ system} - E^{low}_{mode\ l\ system} \quad (1.47)$$

The *real* system represents the full molecular geometry including all atoms and it is treated using a low-level of theory. The *model* system contains the part of the system that is treated at the high level. Superscripts "high" and "low" mean high- and low-level of calculations used in the ONIOM method. Both high and low level calculations need to be carried out for the model system. The scheme can be extended to three or multiple layers. The ONIOM method is so flexible that the final choice of model combination is left to the user and has been found very successful in the theoretical studies of large and complicated systems.

1.14 Objectives of the Present Work

The objectives of the present Ph. D. work are:

1. Density functional theory (DFT) calculations of electronic properties of DNA bases responsible for its damage.
2. Theoretical studies on platinum complexes to find out their reactive nature.
3. Quantum chemical calculations on the anticancer drug molecules to determine their structure-activity/property relationship.
4. To investigate the binding mechanism of a platinum anticancer drug, *cis*-amminedichloro(2-methylpyridine)platinum(II) (AMD473)
5. Interaction of AMD473 with DNA using DFT and hybrid QM/MM methods.

References

(1) Avery, O.T.; Macleod, C.; McCarthy, M. Studies on the chemical nature of the substance including transformation of pneumococcal types. *J. Exp. Med.* **79**, 137–157 (1944)

(2) Chargaff, E. Some recent studies on the composition and structure of nucleic acids. *J. Cell. Physiol. Suppl.* **38**, 41–49 (1951)

(3) Chargaff, E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**, 201–209 (1950)

(4) Watson J.D.; Crick, F.H.C. A structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953)

(5) Wang, A.H.J. *et al* Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **282**, 680–686 (1979)

(6) Zhang, R.B.; Zhang, K.; Eriksson, L.A. Theoretical studies of damage to 3'-uridine monophosphate induced by electron attachment. *Chem. Eur. J.* **14**, 2850–2856 (2008)

(7) Schyman, P.; Laaksonen, A. On the effect of low-energy electron induced DNA strand break in aqueous solution: A theoretical study indicating guanine as a weak link in DNA. *J. Am. Chem. Soc.* **130**, 12254–12255 (2008)

(8) www.who.int, *World Health Organization* (2005)

(9) www. Cancer.org

(10) Pardee, B.; Stein, G.S. *The Biology and Treatment of Cancer: Understanding Cancer* (Wiley-Blackwell, UK, 2009)

(11) Peyrone, M. On the influence of ammonia on platinum chloride. *Ann. Chem. Pharmacol.* **51**, 1–29 (1845)

(12) Rosenberg, B.; Van Camp, L.; Krigas, T. Inhibition of cell division in Escherichia Coli by electrolysis products from a platinum electrode. *Nature* **205**, 698–699 (1965)

(13) Rosenberg, B.; Van Camp, L.; Trosko, J.E.; Mansour, V.H. Platinum compounds: A new class of potent antitumour agents. *Nature* **222**, 385–386 (1969)

(14) Reedijk, J. New clues for platinum antitumor chemistry: Kinetically controlled metal binding to DNA. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3611–3616 (2003)

(15) Giaccone, G. Clinical perspective on platinum resistance. *Drugs* **59**, 9–17 (2000)

(16) Rose, P.G. *et al* Concurrent cisplatin-based radiotherapy and chemotherapy for locally advanced cervical cancer. *N. Engl. J. Med.* **340**, 1144–1153 (1999)

(17) Haxton, K.J.; Burt, H.M. Hyperbranched polymers for controlled release of cisplatin. *Dalton Trans.* **43**, 5872–5875 (2008)

(18) Malik, N.; Evagorou, E.G.; Duncan, R. A novel approach to cancer chemotherapy. *Anticancer Drugs* **10**, 767–776 (1999)

(19) Jamieson, E.R.; Lippard, S.J. Structure, recognition, and processing of cisplatin DNA-adducts. *Chem. Rev.* **99**, 2467–2498 (1999)

(20) Gately, D.P.; Howell, S.B. Cellular accumulation of the anticancer agent cisplatin: a review. *Br. J. Cancer* 67, 1171–1176 (1993)

(21) Andrews, P.A. *In Platinum-Based Drugs in Cancer Therapy* (Kelland, L.R., Farrell, N.P., Eds., Humana Press, Totowa, 2000)

(22) Lin, X.J.; Okuda, T.; Holzer, A.; Howell, S.B. The copper transporter CTR1 regulates cisplatin uptake in saccharomyces cerevisiae. *Mol. Pharmacol.* **62**, 1154–1159 (2002)

(23) Ishida, S.; Lee, J.; Thiele, D.J.; Herskowitz, I. Uptake of the anticancer drug cisplatin mediated by the copper transporter Ctr1 in yeast and mammals. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14298–14302 (2002)

(24) Katano, K *et al* Acquisition of resistance to cisplatin is accompanied by changes in the cellular pharmacology of copper. *Cancer Res.* **62**, 6559–6565 (2002)

(25) Payet, D.; Gaucheron, F.; Sip, M.; Leng, M. Instability of the monofunctional adducts in cis-[Pt(NH3)2(N7-N-methyl-2-diazapyrenium)Cl](2+)-modified DNA: rates of cross-linking reactions in cis-platinum-modified DNA. *Nucleic Acids Res.* **21**, 5846–5851 (1993)

(26) Naser, L.J.; Pinto, A.L.; Lippard, S.J.; Essigmann, J.M., Chemical and biological studies of the major DNA adduct of *cis*-diamminedichloroplatinum(II), *cis*-[$Pt(NH_3)_2${d(GpG)}], built into a specific site in a viral genome. *Biochemistry* **27**, 4357-4367 (1988)

(27) Poklar, N. *et al* Influence of cisplatin intrastrand crosslinking on the-conformation, thermal stability, and energetics of a 20-mer DNA duplex. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 7606-7611 (1996)

(28) Hayes, D.; Cvitkovic, E.; Golbey, R.; Scheiner, E.; Krakoff, I.H. Amelioration of renal toxicity of high dose cis-platinum diammine dichloride (CPDD) by mannitol induced diuresis. *Proc. Am. Assoc. Cancer Res.* **17**, 169–178 (1976)

(29) Cleare, M.J.H.; Hoeschele, J.D. Anti-tumour platinum compounds relationship between structure and activity. *Plat. Met. Rev.* **17**, 2–13 (1973)

(30) Cleare, M.J.H.; Hoeschele, J.D. Studies on the antitumor activity of group VIII transition metal complexes. Part I. Platinum (II) complexes. *Bioinorg. Chem.* **2**, 187–210 (1973)

(31) Lebwohl, D.; Canetta, R. Clinical development of platinum complexes in cancer therapy: an historical perspective and an update. *Eur. J. Cancer* **34**, 1522–1534 (1998)

(32) Uchida, N. *et al* Antitumor efficacy of nedaplatin, a novel platinum complex, with cyclophosphamide in murine and human tumor model. *Anticancer Res.* **18**, 3375–3379 (1998)

(33) Misset, J.L., Oxaliplatin in practice. *Br. J. Cancer* **77**, 4–7 (1998)

(34) Lebwohl, D.; Canetta, R., Clinical development of platinum complexes in cancer therapy: an historical perspective and an update. *Eur. J. Cancer* **34**, 1522–1534 (1998)

(35) Boulikas, T.; Vougiouka, M. Cisplatin and platinum drugs at the molecular level. *Oncol. Rep.* **10**, 1663–1682 (2003)

(36) Uchida, N. *et al* Sequence-dependent antitumour efficacy of combination chemotherapy of nedaplatin, a novel platinum complex, with 5-fluorouracil in an in vivo murine tumour model. *Eur. J. Cancer* **34**, 1796–1801 (1998)

(37) Kostova, I. Platinum complexes as anticancer agents. *Recent Patents Anti-Cancer Drug Discov.* **1**, 1–22 (2006)

(38) Wang, D.; Lippard, S.J. Cellular processing of platinum anticancer drugs. *Nat. Rev. Drug. Discovery.* **4**, 307–320 (2005)

(39) Reedijk, J. Why does cisplatin reach guanine-N7 with competing S-donor ligands available in the cell? *Chem. Rev.* **99**, 2499–2510 (1999)

(40) Chen, Y.; Guo, Z.J.; Parsons, S.; Sadler, P.J. Stereospecific and kinetic control over the hydrolysis of a sterically hindered platinum picoline anticancer complex. *Chem. Eur. J.* **4**, 672–676 (1998)

(41) Holford, J.; Sharp, S.Y.; Murrer, B.A.; Abrams, M.; Kelland, L.R. In vitro circumvention of cisplatin resistance by the novel sterically hindered platinum complex AMD473. *Br. J. Cancer* 77, 366–373 (1998)

(42) Kelland, L.R. *et al* Discovery and development of platinum complexes designed to circumvent cisplatin resistance. *J. Inorg. Biochem.* 77, 111–115 (1999)

(43) Munk, V.P. *et al* Investigations into the interactions between DNA and conformationally constrained pyridylamineplatinum(II) analogues of AMD473 *Inorg. Chem.* **42**, 3582–3590 (2003)

(44) Farrell, N. *et al* Cytostatic trans-platinum(II) complexes. *J. Med. Chem.* **32**, 2240–2241 (1989)

(45) Perez, J.M.; Fuertes, M.A.; Alonso, C.; Navarro-Ranninger, C. Current status of the development of trans-platinum antitumor drugs. *Crit. Rev. Oncol. Hematol.* **35**, 109–120 (2000)

(46) Treskes, M.; Vijgh, W.J.F. WR2721 as a modulator of cisplatin-and carboplatin-induced side effects in comparison with other chemoprotective agents: a molecular approach. *Cancer Chemother. Pharmacol.* **33**, 93–106 (1993)

(47) Kelland, L.R. *et al* Ammine/amine platinum(IV) dicarboxylates: a novel class of platinum complex exhibiting selective cytotoxicity to intrinsically cisplatin-resistant human ovarian carcinoma cell lines. *Cancer Res.* **52**, 822–828 (1992)

(48) Fokkema, E.; Groen, H.J.; Helder, M.N.; De Vries, E.G.; Meijer, C. JM216-, JM118-, and cisplatin-induced cytotoxicity in relation to platinum-DNA adduct formation, glutathione levels and p53 status in human tumour cell lines with different sensitivities to cisplatin. *Biochem. Pharmacol.* **63**, 1989–1996 (2002)

(49) Farrell, N. *Metal ions in Biological Systems*. (Marcel Dekker, New York, Basel, 2004)

(50) Komeda, S.; Lutz, M.; Spek, A.L.; Chikuma, M.; Reedijk, J. New antitumor-active azole-bridged dinuclear platinum(II) complexes: synthesis,

characterization, crystal structures, and cytotoxic studies. *Inorg. Chem.* **39**, 4230–4236 (2000)

(51) Wheate, N.J.; Collins, J.G. Multi-nuclear platinum complexes as anti-cancer agents. *Coord. Chem. Rev.* **241**, 133–145 (2003)

(52) Zou, Y.; Van Houten, B.; Farrell, N. Sequence specificity of DNA-DNA interstrand cross-link formation by cisplatin and dinuclear platinum complexes. *Biochemistry* **33**, 5404 –5410 (1994)

(53) Fokkema, E.; Groen, H.J.; Helder, M.N.; De Vries, E.G.; Meijer, C. JM216-, JM118-, and cisplatin-induced cytotoxicity in relation to platinum-DNA adduct formation, glutathione levels and p53 status in human tumour cell lines with different sensitivities to cisplatin. *Biochem. Pharmacol.* **63**, 1989–1996 (2002)

(54) Andrews, P.A.; Howell, S.B. Cellular pharmacology of cisplatin: perspectives on mechanisms of acquired resistance. *Cancer Cell.* **2**, 35–43 (1990)

(55) Andrews, P.A.; Schreffer, M.A.; Murphy, M.P.; Howell, S.B. Enhanced potentiation of cisplatin cytotoxicity in human ovarian carcinoma cells by prolonged glutathione depletion. *Chem. Biol. Interact.* **65**, 51–58 (1988)

(56) Mellish, K.J.; Kelland, L.R.; Harrap, K.R. In vitro platinum drug chemosensitivity of intrinsic and acquired cisplatin-resistant human cervical carcinoma cell lines. *Br. J. Cancer* **68**, 240–250 (1993)

(57) Dixit, K.S.; Mitra, S.N. Bioinformatics in drug discovery. *Curr. Res. Inf. Pharm. Sci.* **3**, 2–6 (2002)

(58) Balzarini, J. *et al* Role of thymidine kinase in the inhibitory activity of 5-substituted-2¢-deoxyuridines on the growth of human and murine tumor cell lines. *Biochem. Pharmacol.* **31**, 1089–95 (1982)

(59) Ray, A.S. *et al* Insights into the molecular mechanism of inhibition and drug resistance for HIV-1 RT with carbovir triphosphate. *Biochemistry* **41**, 5150–5162 (2002)

(60) Hanson, R.L. *et al* Regioselective enzymatic aminoacylation of lobucavir to give an intermediate for lobucavir prodrug. *Bioorg. Med. Chem.* **8**, 2681–2687 (2000)

(61) Groove, C.L. *et al* Anticancer activity of beta-L-dioxolane-cytidine, a novel nucleoside analogue with the unnatural L configuration. *Cancer Res.* **55**, 3008–3011 (1995)

(62) Jeong, L.S. *et al* Asymmetric synthesis and biological evaluation of beta-L-(2R,5S)-and alpha-L-(2R,5R)-1,3-oxathiolane-pyrimidine and purine nucleosides as potential anti-HIV agents. *J. Med. Chem.* **36**, 181–195 (1993)

(63) Clare, B.W.; Supuran, C.T. A perspective on quantitative structureactivity relationships and carbonic anhydrase inhibitors. *Expert. Opin. Drug. Metab. Toxicol.* **2**, 113–137 (2006)

(64) Verma, R.P. Understanding apoptosis in terms of QSAR. *Anticancer Agents Med. Chem.* **6**, 41–52 (2006)

(65) Kozikowski, A.P.; Roth, B.; Tropsha, A. Why academic drug discovery makes sense. *Science* **313**, 1235–1236 (2006)

(66) Huang, R.; Wallqvist, A.; Covell, D.G. Assessment of in vitro and in vivo activities in the National Cancer Institute's anticancer screen with respect to chemical structure, target specificity, and mechanism of action. *J. Med. Chem.* **49**, 1964–1979 (2006)

(67) Ren, S.S.; Lien, E.J. Anticancer agents: tumor cell growth inhibitory activity and binary QSAR analysis. *Curr. Pharm. Des.* **10**, 1399–1415 (2004)

(68) Martin, Y.C. *Quantitative Drug Design: A Critical Introduction* (Marcel Dekker, New York, 1978)

(69) Richards, W.G. *Quantum Pharmacology* (Butterworths, New York, 1983)

(70) Franke, R. *Theoretical Drug Design Methods* (Elsevier, Amsterdam, 1984)

(71) Schleyer, P.V.R. *et al Encyclopedia of Computational Chemistry* (Wiley & Sons, UK, 1998)

(72) Finn, P.W. Kavraki, L.E. Computational approaches to drug design. *Algorithmica*, **25**, 347–371 (1999)

(73) Koch, W.; Holthausen, M.C. *A Chemist's Guide to Density Functional Theory* (Wiley-VCH, Germany, 2000)

(74) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev. B* **135**, 864–871 (1964)

(75) Kohn, W.; Sham, L.J. Self-consistent equations including exchange and correlation effects. *Phys. Rev. A* **140**, 1133–1138 (1965)

(76) Langreth D.C.; Perdew J.P. The exchange-correlation energy of a metallic surface. *Solid State Comm.* **17**, 1425–1429 (1975)

(77) Perdew J.P. Density-functional approximation for the correlation energy of the

inhomogeneous electron gas. *Phys. Rev. B* **33**, 8822–8824 (1986)

(78) Becke A.D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988)

(79) Lee C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988)

(80) Sulpizi, L.; Folkers, G.; Rothlisberger, U.; Carloni P.; Scapozza L. Applications of density functional theory-based methods in medicinal chemistry. *Quant. Struct.-Act. Rel.* **21**, 173–181 (2002)

(81) Carloni, P.; Rothlisberger, U.; Parrinello, M. The role and perspective of ab-initio molecular dynamics in the study of biological systems. *Acc. Chem. Res.* **35**, 455–464 (2002)

(82) Becke, A.D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372–1377 (1993)

(83) Iczkowski, R.P.; Margrave, J.L. Electronegativity. *J. Am. Chem. Soc.* **83**, 3547–3551 (1961)

(84) Parr R.G.; Pearson R.G. Absolute hardness: companion parameter to absolute electronegativity. *J. Am. Chem. Soc.* **105**, 7512–7516 (1983)

(85) Koopmans, T.A. Ordering of Wave functions and eigenvalues to the individual electrons of an atom. *Physica* **1**, 104–113 (1933)

(86) Parr, R.G.; Szentpaly, L.V.; Liu, S. Electrophilicity Index. *J. Am. Chem. Soc.* **121**, 1922–1924 (1999)

(87) Mendez, F.; Gazquez, J.L. Chemical reactivity of enolate Ions: The local hard and soft acids and bases principle viewpoint. *J. Am. Chem. Soc.* **116**, 9298–9301 (1994)

(88) Yang, W.; Mortier, W.J. The use of global and local molecular parameters for the analysis of the gas-phase basicity of amines. *J. Am. Chem. Soc.* **108**, 5708–5711 (1986)

(89) Chattaraj, P.K.; Maiti, B.; Sarkar, U. Philicity: A unified treatment of chemical reactivity and selectivity. *J. Phys. Chem. A* **107**, 4973–4975 (2003)

(90) Hansch, C. *et al* The correlation of biological activity of plant growth regulators and chloromycetin derivatives with Hammett constants and partition coefficients. *J. Am. Chem. Soc.* **85**, 2817–2824 (1963)

(91) Hansch, C. Fujita, T. ρ-σ-π Analysis: method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.* **86**, 1616–1626 (1964)

(92) Debnath, A.K. Quantitative structure-activity relationship (QSAR) paradigm –Hansch era to new millennium. *Mini. Rev. Med. Chem.* **1**, 187–195 (2001)

(93) Myers, R.H. *Classical and Modern Regression with Applications.* (Duxbury Press, Belmont, Calif. 1990)

(94) Roy, D.R.; Giri, S.; Chattaraj, P.K. Arsenic Toxicity: An atom counting and electrophilicity based protocol. *Mol. Div.* **13**, 551–556 (2009)

(95) Roy, D.R. *et al* Analyzing toxicity through electrophilicity. *Mol. Div.* **10**, 119–131 (2006)

(96) Roy, D.R.; Subramanian, V.; Chattaraj, P.K. An electrophilicity based analysis of toxicity of aromatic compounds towards tetrahymena pyriformis. *QSAR Comb. Sci.* **25**, 114–122 (2006)

(97) Singh, U.C.; Kollman, P.A. A combined *ab initio* quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the $CH_3Cl + Cl^-$ exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.* **7**, 718–730 (1986)

(98) Field, M.J.; Bash, P.A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **11**, 700–733 (1990)

(99) Thompson, M.A. QM/MMpol: A consistent model for solute/solvent polarization. Application to the aqueous solvation and spectroscopy of formaldehyde, acetaldehyde, and acetone. *J. Phys. Chem.* **100**, 14492–14507 (1996)

(100) Born, M.Z. Volumes and hydration warmth of ions. *Physik* **1**, 45–48 (1920)

(101) Onsager, L. Electric moments of molecules in liquids. *J. Am. Chem. Soc.* **58**, 1486–1493 (1936)

(102) Kirkwood, J.G. The dielectric polarization of polar liquids. *J. Chem. Phys.* **7**, 911–919 (1939)

(103) Miertus, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **55**, 117–129 (1981)

(104) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **105**, 2999–3093 (2005)

(105) Chipman, D.M. Reaction field treatment of charge penetration. *J. Chem. Phys.* **112**, 5558–5565 (2000)

(106) Andzelm, J.; Kolmel, C.; Klamt, A. Incorporation of solvent effects into density functional calculations of molecular energies and geometries. *J. Chem. Phys.* **103**, 9312–9320 (1995)

(107) Truong, T.N.; Stefanovich, E.V. Analytical first and second energy derivatives of the generalized conductor like screening model for free energy of solvation. *J. Chem. Phys.* **103**, 3709–3717 (1995)

(108) Dapprich, S.; Komaromi, I.; Byun, K.S.; Morokuma, K.; Frisch, M.J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Struct.(THEOCHEM)* **1**, 461–462 (1999)

(109) Svensson, M.; Humbel, S.; Morokuma, K. Energetics using the single point IMOMO (integrated molecular orbital+molecular orbital) calculations: Choices of computational levels and model system. *J. Chem. Phys.* **105**, 3654–3661 (1996)

(110) Karadakov, P.B.; Morokuma, K. ONIOM as an efficient tool for calculating NMR chemical shielding constants in large molecules. *Chem. Phys. Lett.* **317**, 589–596 (2000)

(111) Vreven, T.; Morokuma, K. On the application of the IMOMO (integrated molecular orbital + molecular orbital) method. *J. Comput. Chem.* **21**, 1419–1432 (2000)

Chapter 2

Investigation of Electron Affinities of DNA and RNA Bases

In this chapter, we have performed a series of DFT calculations on the nucleobases to determine their electron affinities that aid our understanding towards the primary radiation processes in DNA. Different exchange-correlation functionals and basis sets are used to explore their influence on electron affinities of thymine, cytosine, adenine, guanine, and uracil bases. Results showed that vertical electron affinities are always negative for all studied level of theory in agreement with the experimental observation. On the other hand, the adiabatic electron affinities depend on both the exchange-correlation potential and basis set quality. Basis set with diffuse functions in combination with B3LYP functional confirms positive value of adiabatic electron affinities of thymine and uracil in consistent with experimental data. We always found an unstable covalently bound anion for adenine as predicted by previous theoretical calculations. The negative adiabatic electron affinity value of cytosine agrees well with the available theoretical results but does not match with the experimental observation. For guanine the adiabatic electron affinity values calculated at B3LYP/6-311++G** level was found to be very high compared to that obtained at B3LYP/6-31G and B3LYP/TZVP levels confirming the formation of mixed covalent-dipole character [Sarmah and Deka, *Mol. Sim.* **34**, 879–7791 (2008)]

2.1 Introduction

The electronic properties of nucleic acid bases are of fundamental importance to gain knowledge regarding the storage and transfer mechanisms of genetic information. When ionizing radiation interacts with nucleic acid bases, low energy electrons are produced that are trapped by nucleobases generating radical anions which can lead to both DNA base damage and strand breakage. In many cases the specific DNA damage may either block replication and transcription or generate mutation by miscoding during replication. Adiabatic electron affinities (AEAs) are the thermodynamic parameters which govern the ease of reduction of each nucleobase. In this context, the determination of electron affinities (EA) of DNA and RNA bases can be of great importance owing to their significance in understanding the DNA radiation damage as well as excess electron transfer through DNA. There have been a series of reports both experimental[1-9] and theoretical,[10-23] concerning the investigation of electronic properties of nucleobases. The vertical electron affinities (VEAs) of DNA and RNA bases have been reported experimentally by Aflatooni *et al.*[5] using electron transmission spectroscopy (ETS) technique where they found vertical attachment energies (VAEs) for all nucleobases are > 0 or in the other sense their vertical electron affinities are negative. In the very early 90s experimental measurement of the AEAs of nucleobases was first estimated and found substantial values with the relative order: Cytosine (C) <Thymine (T) <Uracil (U) <Adenine (A) <<Guanine (G).[1]

In contrast, many computational studies on nucleic acid base anions found negative valence adiabatic electron affinities for some of the nucleobases. For example PMP2 calculations with the 6-311+G(2df,p) basis set yields negative AEAs values for all DNA bases in the range -0.12 to -0.73 eV.[19] In the early 90s, noting the large dipole moments of nucleic acid bases (~5D), Oyler and Adamowicz[15] conducted calculations which found stable anions of uracil in which their electrons were bound via dipole interactions often referred as "dipole-bound anion". In covalent anions the extra electron enters to the LUMO of the molecule whereas in dipole-bound anion the excess electron is bound by the dipole fields of the neutral molecule without influencing the structure of the molecule. Following these theoretical predictions two complementary experiments on gas phase nucleic acids base anions were conducted by Desfrancois *et al.*[2] and Bowen and coworkers[3] with

the help of Rydberg electron transfer spectroscopic studies and photoelectron spectroscopy, respectively. Both experiments found stable dipole-bound anions, confirming the theoretical predictions of Oyler and Adamowicz.[15] After verifying the existence of dipole-bound anion, Desfrancois *et al.*[4] estimated simultaneous existence of both valence and dipole-bound states of uracil anion by using Rydberg electron transfer spectroscopy to gas-phase isolated uracil molecules and mixed uracil-argon clusters.

By noting the fact that nucleobases in the gas-phase can form dipole-bound anion and on the other hand in the condensed phase nucleobase anions are conventional (covalent) anions, Hendricks *et al.*[6] performed a series of negative ion photoelectron spectroscopic experiments in gas-phase and solvated uracil cluster anions to assess the point at which the uracil anion converts from dipole-bound to covalently bound state. They observed a sharp peak in the photoelectron spectrum of uracil anion and explained this behavior by suggesting that the uracil anion essentially has the same structure as the neutral which is typical for dipole bound anion. They have also found the vertical detachment energy (VDE) of uracil anion to be 0.093±0.007 eV which is in good agreement with the adiabatic dipole bound EA estimated by Oyler and Adamowicz.[15] The conversion of dipole-bound state to the covalent anion state was observed by the inclusion of a single water molecule indicating a broad photoelectron spectrum of U-(H_2O). They saw a dramatic effect of having both dipole-bound and covalent like features of the complex by the replacement of water molecule by Xe as solvating agent. Similar to the studies of Hendricks and co-workers, Schiedt *et al.*[7] studied the photoelectron spectra of uracil, thymine and cytosine in free and microsolvated form. The reported extrapolated electron affinities from this microsolvation experiments are 0.12, 0.13, and 0.15 eV for T, C, and U respectively.

Density functional theory (DFT) was first used in the determination of AEA for uracil by Desfrancois *et al.*[4] and found a small but positive value of 0.07 eV. Recently, several studies of nucleobases' AEAs have included various DFT computations.[19,20,22,23] Wesolowski *et al.*[20] used double-ξ plus polarization plus diffuse (DZP++) basis set with five different density functionals in order to bracket the true AEAs of nucleobases. They have concluded that U and T covalent anions are bound by ca. 0.05-0.25 eV whereas adenine does not have a stable covalently

bound anion in the gas phase. Although various experimental and theoretical works have been carried out on the nucleobases, true values of AEAs of the bases is still a matter of controversy and the precise determination of their electron affinities remains challenging.

In this chapter, the tautomers of nucleic acid bases shown in Figure 2.1 are chosen on the basis of their stability and biological importance and electron affinities of these bases are determined using hybrid potential B3LYP with different basis sets in order to establish the influence of orbital basis sets in the study of the electronic properties of these biomolecules. The EAs of the selected systems are also calculated using the better predicted basis set in combination with GGA exchange-correlation functionals, PBEOP, PBELYP, and PBEVWN.

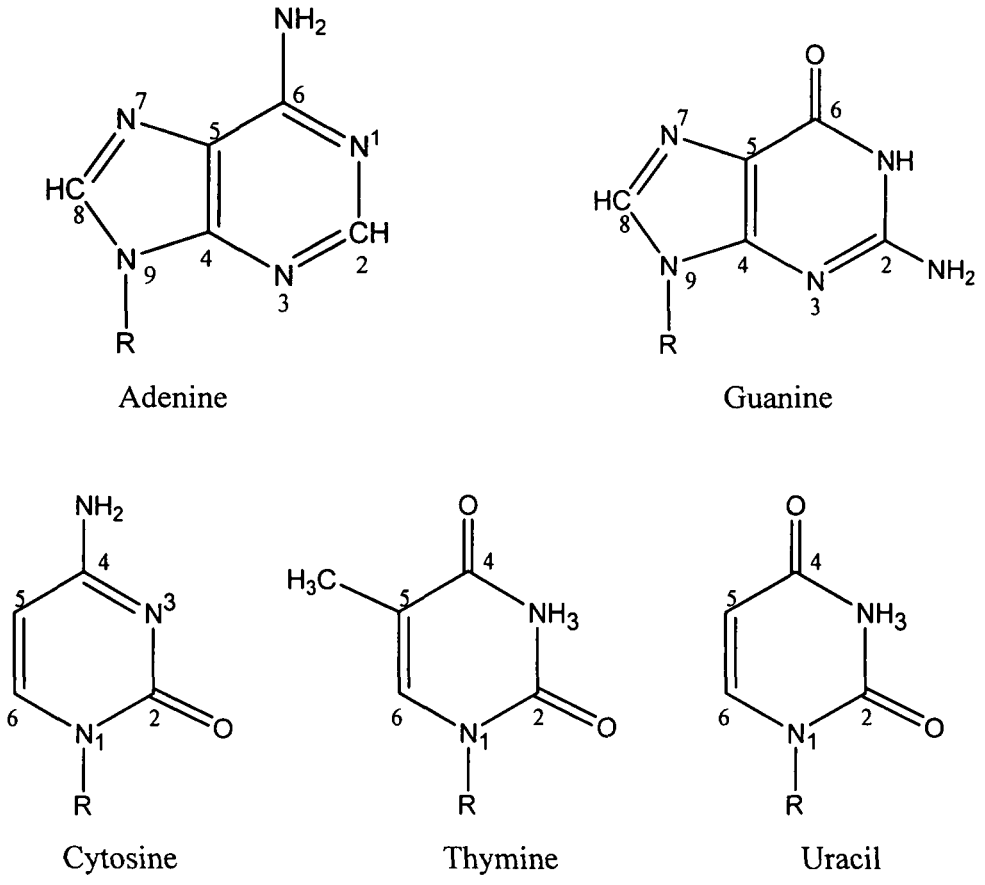


Figure 2.1: Major tautomeric structures of DNA and RNA bases.

2.2 Computational Details

The nucleobases were optimized using DFT theory with hybrid generalized gradient approximation (GGA) exchange-correlation density functional B3LYP in connection with 6-31G, TZVP and 6-311++G** basis sets. B3LYP is Becke's three parameter exchange functional (B3)[24] in combination with Lee, Yang, and Parr (LYP)[25] correlation functional. GGA exchange-correlation density functionals PBEOP, PBELYP and GGE exchange-correlation functional PBEVWN were used in connection with better predicted basis set 6-311++G** to notice their influence in determining the electron affinities of the nucleobases. These are Perdew-Burke-Ernzerhof exchange[26] + one-parameter progressive (OP)[27] correlation (PBEOP), Lee, Yang, and Parr (LYP) correlation (PBELYP) and Vosko, Wilk, Nusair (VWN5) correlation.[28] GGE (generalized gradient exchange) approach was proposed by Hertwig and Koch[29] in which functionals are the combination of GGA exchange functionals and local correlation functionals. All computations were carried out with the GAMESS[30] program package.

The electron affinity is the energy of the neutral molecule minus that of the anion radical

$$EA = E_{neutral} - E_{anion}$$

The adiabatic electron affinity is the difference in total energies between the optimized structures of neutral and anionic systems. The calculation of vertical electron affinity employs the optimized geometry of neutral forms to compute the energy of the corresponding anions.

2.3 Results and Discussion

The adiabatic electron affinities without zero point energy correction, obtained at B3LYP level with 6-31G, TZVP and 6-311++G** basis sets are given in Table 2.1. Although zero-point vibrational energy corrections are not included to our AEA values, we note significant nuclear rearrangement of the anions relative to their neutral counterparts.

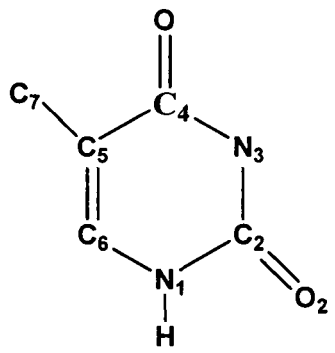
Table 2.1: ZPE uncorrected adiabatic electron affinities (eV) of DNA and RNA bases by different theoretical and experimental methods.

Methods	Uracil	Thymine	Cytosine	Guanine	Adenine
B3LYP/6-31G	-0.56	-0.61	-0.96	-1.13	-1.25
B3LYP/TZVP	-0.19	-0.21	-0.50	-0.79	-0.89
B3LYP/6-311++G**	0.23	0.16	-0.16	-0.14	-0.41
B3LYP/DZP++[a]	0.24	0.20	0.03	-0.01	-0.28
MP2/6-31+G(d)// 6-31G*[b]	-0.51	-0.54	-0.83	-1.79	-1.47
Expt[c]	0.054±0.035	0.068±0.020	—	—	0.012±0.005
Expt[d]	0.150±0.120	0.120±0.120	0.130±0.120	—	—

a) Ref. 20 (ZPVE corrected)
b) Ref 10
c) Ref. 2
d) Ref. 7 (extrapolated values)

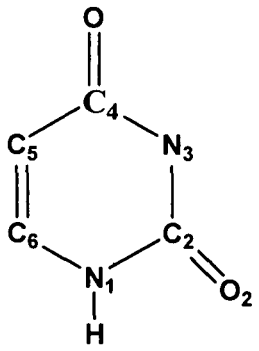
The AEA values for 6-31G and TZVP basis sets (without diffuse functions) are negative for all bases indicating unstable anion formation whereas for larger basis set containing diffuse functions (6-311++G**) give positive EA values for U and T in agreement with some of the previous experimental[2,4,7] and theoretical results.[10,15] This is because anions are better predicted with diffuse functions as the description of the spatial expansion of electron density for anions needs additional diffuse basis functions. The attachment of an electron to the planar neutral molecule leads to the deviation of the system from planarity. The anions of all nucleobases deviate largely from planarity which is evident from the change in dihedral angles of the anions. The main torsional angles of thymine and uracil anions obtained at the

B3LYP/6-311++G** level are shown in Figure 2.2. The same dihedral angles of the corresponding neutral systems are equal to 0 or 180 degrees.



Thymine anion

$N_1C_2N_3C_4 = 0.914$
$C_2N_3C_4O_4 = 176.63$
$O_4C_4C_5C_6 = 176.83$
$C_4C_5C_6N_1 = 11.17$
$C_5C_6N_1C_2 = -15.64$
$C_6N_1C_2N_3 = 9.45$
$O_2C_2N_3C_4 = -179.10$
$N_1C_6C_5C_7 = -172.24$



Uracil anion

$N_1C_2N_3C_4 = -2.72$
$C_2N_3C_4O_4 = 177.63$
$O_4C_4C_1C_6 = 176.51$
$C_4C_5C_6N_1 = 10.20$
$C_5C_6N_1C_2 = -21.83$
$C_6N_1C_2N_3 = 12.71$
$O_2C_2N_3C_4 = 178.32$

Figure 2.2: Torsional angle values (in degrees) for thymine and uracil anions at the optimized geometries. The same torsional angles for neutral system are equal to 180 or 0 degree.

The variation in adiabatic electron affinity with respect to basis set is displayed graphically in Figure 2.3. We observed similar trends of AEA values using 6-31G and TZVP basis sets but they differ significantly when diffuse functions are included. The trend obtained by using 6-311++G** basis set is shifted to higher EA values from that of the other two basis sets. Almost similar trends are observed for U, T, and C at three basis sets. The EA value obtained by basis set with diffuse function is slightly higher than the regular trend predicted with other basis sets for A and much more for G.

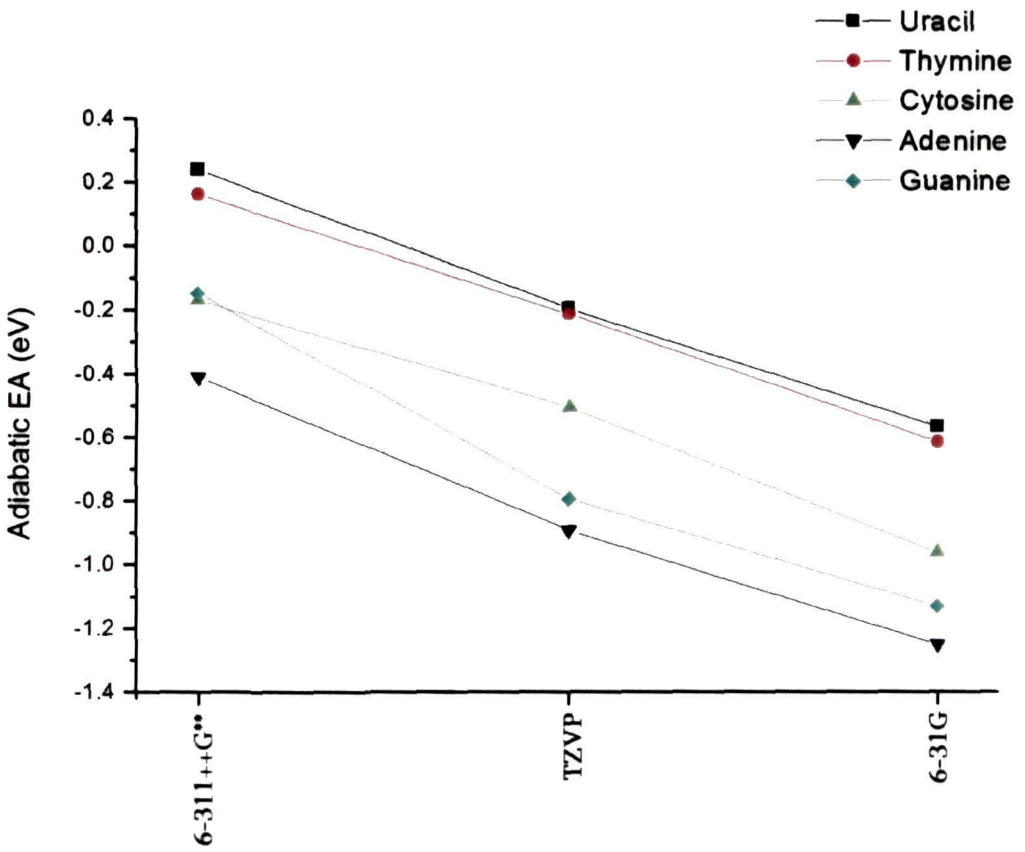


Figure 2.3: Variation of the DFT-B3LYP calculated adiabatic electron affinities of the DNA/RNA bases at different basis sets. The regular trend lines differ significantly for guanine when diffuse functions are included.

The electron affinities of purine bases are very sensitive to diffuse functions. This sensitivity is high for guanine. Except for the concordance in sign, its AEA value with 6-311++G** basis set is much higher than those obtained by Sevilla *et al.*[10] This discrepancy could be explained by considering the fact that their MP2 results arise from single-point computations on geometries optimized at the HF level. The dipole moment of G is sufficiently high (6.87 D at B3LYP/6-311++G** level) and the higher EA value for G at this level is probably due to the dipole bound contamination. All levels of theory give most negative value for adenine in agreement with earlier DFT calculations.[20,22] The dipole moment of adenine at B3LYP/6-311++G** level is found to be 2.46 D which is similar with that of Russo *et al.*[22] This amount of dipole moment should not be enough to bind the electrons via dipolar interactions.[22,31,32] Distortion of the structure of its anion is also not so

much influencing. Thus adenine does not have a stable covalently bound anion or dipole bound anion in the gas phase. The reason for getting slightly higher trend at 6-311++G** basis set as compared to the other basis set remains unclear. Li *et al.*[23] reported that the purine anions exhibit a mixed covalent-dipole character when large, diffuse basis sets are employed. Under biological conditions DNA bases are not isolated but rather solvated by water molecules. Since the dipole-bound states are not believed to be relevant to aqueous systems,[10] conventional valence electron affinities are the only biological relevant quantity. Thus the covalent (negative) electron affinities of purine bases are perhaps best estimated using small basis sets that constrain the electron density on the molecular framework.[33]

Similar trends are observed for the adiabatic electron affinities of pyrimidine bases with three basis sets. Experimental studies for AEA of U give positive value ranging from 0.030 to 0.093 eV.[2,4,7] Although the experimental determinations of EAs for U and T by Desfrancois *et al.*[2] and Schiedt *et al.*[7] yielded positive EAs, their values are significantly different (Table 2.1). Desfrancois *et al.*[2] reported the dipole-bound EAs values of the nucleobases whereas Schiedt *et al.*[7] estimated the valence EAs of the pyrimidine bases by extrapolation of their data for microsolvated anions that may be the reason for disparity between these two experimental results. Our computations yield negative EA for U with 6-31G and TZVP basis sets. This is probably due to the lack of diffuse functions that are mandatory in the description of anions. The AEA value of U obtained with 6-311++G** (0.239 eV) is in good agreement with those determined by Schiedt *et.al.*[7] The employed computational level gives positive AEA with significant geometry distortion of the anions with respect to their neutral form. This suggests the entrance of the extra electron in the LUMO of the neutral molecule. The dipole moment of uracil is 4.19 D at B3LYP/6-311++G**, and an additional contribution to the stability can arise from the dipole-electron interaction. Our B3LYP/6-311++G** value is very similar to the ZPE corrected value at B3LYP/DZP++ level.[20] Similarly the sign of adiabatic electron affinity values of thymine for basis sets without diffuse functions are inconsistent with the previously reported experimental results.[2,7] This is again due to the lack of diffuse functions which can describe the spatial expansion of the electron density of the anion. The positive EA value obtained with 6-311++G** basis set is in agreement with experimental and most of the theoretical data[2,7,10,19,20,22,23] data. This suggests a covalent electron attachment that can be confirmed by the rather large

rearrangement of the geometry in the anion (Figure 2.2). But the dipole bound state can also be possible due to dipole moment value that is 4.54D (B3LYP/6-311++G**). The AEA value obtained with 6-311++G** basis set falls in the range reported by Schiedt *et al.*[7] For cytosine we found negative adiabatic electron affinities at all level of calculations that agree with some of the previous theoretical[19,22,23] value but inconsistent with the experimental information available in the literature.[7,9] The AEAs obtained by Wesolowski *et al.*[20] oscillate between small positive and negative values for the three most reliable functional combinations and they have concluded that the AEA of cytosine remains ambiguous. Due to the large dipole moment of this molecule (6.77 D at B3LYP/6-311+G** level) the contributions from dipole-bound state must be considered. At this level we found comparatively higher electron affinity value than with the other two basis sets but disagreement in sing with the experimental data prevents a definitive conclusion. Table 2.2 lists the calculated vertical electron affinities (VEA) values. Some previous theoretical and experimental results are also reported for the purpose of comparison.

Table 2.2: Zero-point uncorrected vertical electron affinities (in eV) of DNA and RNA bases by different theoretical and experimental methods.

Methods	Uracil	Thymine	Cytosine	Adenine	Guanine
B3LYP/6-31G	-0.94	-1.02	-1.32	-1.51	-1.86
B3LYP/TZVP	-0.59	-0.64	-0.87	-1.08	-1.58
B3LYP/6-311++G**	-0.23	-0.26	-0.25	-0.41	-0.15
MP2/6-31+G(D)[a]	-0.19	-0.32	-0.40	-0.74	-1.23
Expt[b]	-0.22	-0.29	-0.32	-0.54	—
Expt[c]	—	0	-0.55	-0.45	—

(a) Ref. 10, scaled results

(b) Ref. 5

(c) Ref. 8

In Figure 2.4 we display the variation of vertical electron affinities with respect to the basis sets. All values are found to be negative in good agreement with the results of ETS technique by Aflatooni *et al.*[5] The trend lines for C and A obtained at B3LYP/6-311++G** level slightly differ and for G it substantially differs at the same level of theory from the other two basis sets. The vertical EAs obtained with 6-31G basis set are much more negative whereas values obtained by B3LYP functional coupled with TZVP basis set are twice as large as those obtained by experimental measurement. The results obtained by B3LYP/6-311++G** method are closer to the experimental counterparts.[5]

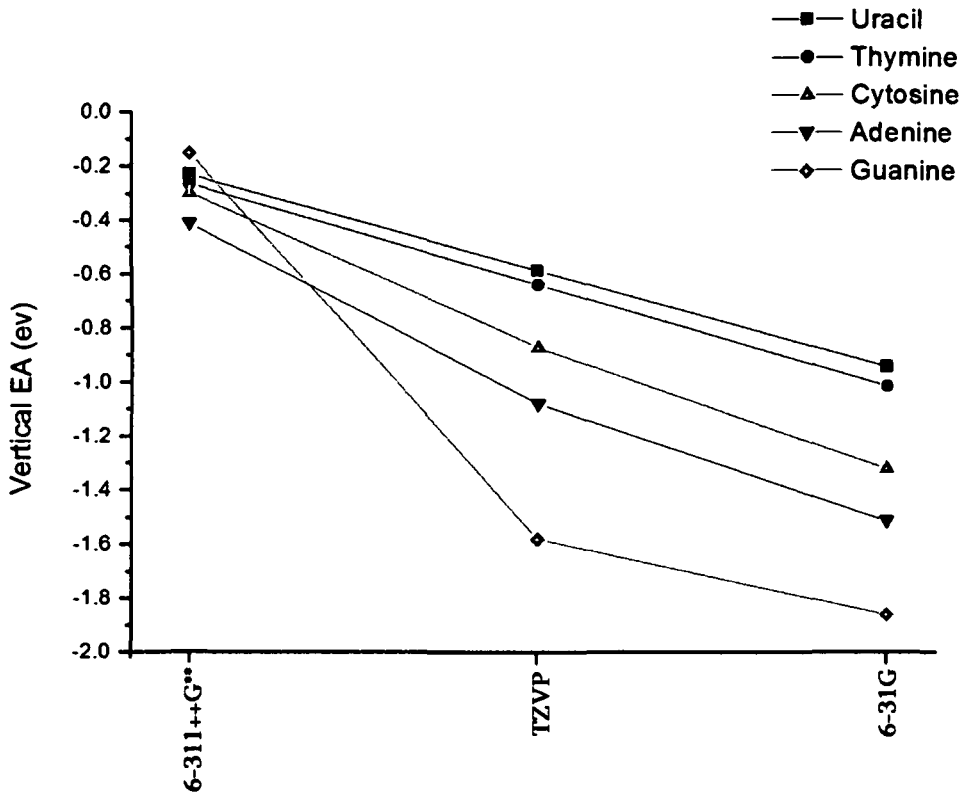


Figure 2.4: Variation of the DFT-B3LYP calculated vertical electron affinities of the DNA/RNA bases at different basis sets.

The adiabatic electron affinities computed with the better predicted basis set 6-311++G** for two GGA exchange correlation functionals PBEOP and PBELYP are given in Table 2.3. Neither the PBEOP nor PBELYP calculations indicate positive AEA as found experimentally. Both these methods yield negative EA for U and T with more negative value for U anion than T which is inconsistent with

reported experimental and theoretical data. However AEA values of cytosine obtained with these two GGA functionals move slightly more close to the previous theoretical[19,22,23] determination than that of hybrid B3LYP functional. But its sign is inconsistent with the experimental information available in the literature.[7,9] Due to the large dipole moment of this molecule (~ 7.0 D) an additional contributions from dipole-bound state must be considered.

Table 2.3: Electron Affinities (in eV, ZPE uncorrected) calculated by PBEOP and PBELYP functionals with 6-311++G** basis set.

	Adiabatic EA		Vertical EA	
System	PBEOP	PBELYP	PBEOP	PBELYP
Uracil	-0.320	-0.182	-0.246	-0.146
Thymine	-0.122	-0.003	-0.250	-0.071
Cytosine	-0.150	-0.129	-0.277	-0.099
Adenine	-0.282	-0.273	-0.331	-0.238
Guanine	-0.053	0.054	-0.062	-0.038

On the other hand, these two functionals predict comparable AEAs for purines. For adenine the AEA is found to be negative with both PBEOP and PBELYP methods in agreement with earlier DFT calculations[20,22] and experimental[8] gas-phase value. The PBELYP/6-311++G** level predicts small positive value of AEA for guanine while small negative value is obtained with PBEOP method. This oscillation is in agreement with the result obtained by Wesolowski *et al.*[20] The AEA values calculated for U, T, and G using PBEVWN functional and 6-311++G** basis set are 0.013 eV, 0.347 eV, and 0.291 eV, respectively. The sign of the adiabatic electron affinities obtained with this functional are in agreement with experiments. But for C and A, we found very high values of AEA using PBEVWN functional.

The PBEOP and PBELYP vertical electron affinities for all nucleobases are listed in Table 2.3. All values are found to be negative in good agreement with the results of ETS technique.[5] However, PBEVWN predicts positive values of vertical

electron affinities for all nucleobases inconsistent with experiment. The variation in VEAs with respect to the PBEOP and PBELYP functionals along with the ETS data of Aflatooni *et al.*[5] are displayed graphically in Figure 2.5.

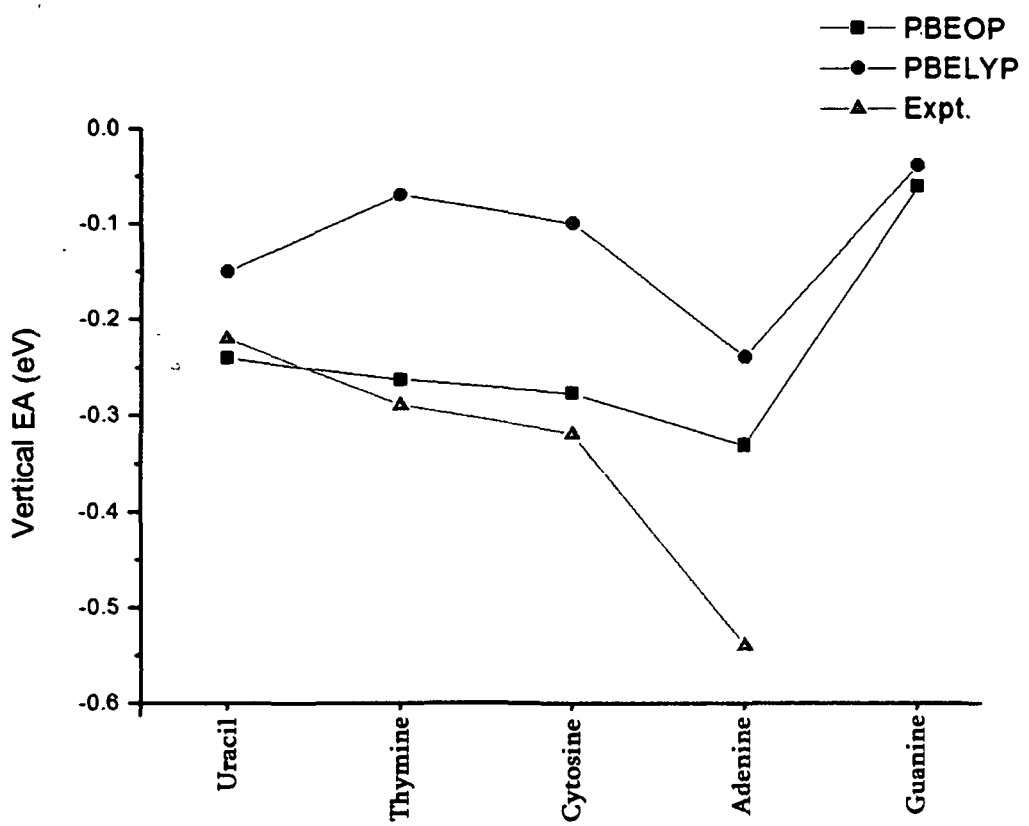


Figure 2.5: Variation of the vertical electron affinities of the DNA/RNA bases.

Although PBELYP predicts higher VEAs for all the selected systems, PBEOP values are in agreement with the experimental values. For U, T, and C the PBEOP values slightly deviate from the experimental one while for A its value deviates more. The small amount of dipole moment of A (2.42D) at this level can not be a reason for this deviation. However, higher value of VEA for G may be due to the additional contribution from dipole bound state with dipole moment 6.88D.

2.4 Conclusions

Computed electron affinities for nucleobases in the gas phase using different basis sets suggest that the VEA of all nucleobases are negative in all cases with

B3LYP, PBEOP, and PBELYP functionals in agreement with available experimental data. The basis sets without diffuse functions give more negative values of VEA as compared to the experimental values. The VEA values for 6-311++G** basis set with B3LYP functional follow the same trend as the other basis sets and close to that of the experimental values. The basis set with diffuse functions confirms the positive value of adiabatic electron affinities of uracil and thymine at B3LYP/6-311++G** level in agreement with the experimental and theoretical results. The dipole moment of cytosine is very high at this level and we can assume dipole bound state for it. For adenine we found an unstable covalently bound anion in the gas phase in agreement with previous theoretical data. Guanine is most sensitive to diffuse functions and its AEA value for 6-311++G** basis set is more than that of available data. This is due to the large dipole moment of the molecule for which it can exhibits mixed covalent-dipole character.

References

(1) Wiley, J.R. *et al* The determination of absolute electron affinities of the purines and pyrimidines in DNA and RNA from reversible reduction potentials. *Biochem. Biophys. Res. Commun.* **180**, 841–845 (1991)

(2) Desfrancois, C.; Abdoul-Carime, H.; Schermann, J.P. Electron attachment to isolated nucleic acid bases. *J. Chem. Phys.* **104**, 7792–7794 (1996)

(3) Hendricks, J.H.; Lyapustina, S.A.; de Clercq, H.L.; Snodgrass, J.T.; Bowen, K.H. Dipole bound nucleic acid base anions studied via negative ion photoelectron spectroscopy. *J. Chem. Phys.* **104**, 7788–7791 (1996)

(4) Desfrancois, C.; Periquet, V.; Bouteiller, Y.; Schermann, J.P. Valence and dipole binding of electrons to uracil. *J. Phys. Chem. A* **102**, 1274–1278 (1998)

(5) Aflatooni, K.; Gallup, G.A.; Burrow, P.D. Electron attachment energies of the DNA bases. *J. Phys. Chem. A* **102**, 6205–6207 (1998)

(6) Hendricks, J.H.; Lyapustina, S.A.; de Clercq, H.L.; Bowen, K.H. The dipole bound-to-covalent anion transformation in uracil. *J. Chem. Phys.* **108**, 8–11 (1998)

(7) Schiedt, J.; Weinkauf, R.; Neumark, D.M.; Schlag, E.W. Anion spectroscopy of uracil, thymine and the amino-oxo and amino-hydroxy tautomers of cytosine and their water clusters. *Chem. Phys.* **239**, 511–524 (1998)

(8) Periquet,V.; Moreau, A.; Carles, S.; Schermann, J.P.; Desfrancois, C. Cluster size effects upon anion solvation of N-heterocyclic molecules and nucleic acid bases. *J. Electron. Spectrosc. Rel. Phenomena.* **106**, 141–151 (2000)

(9) Chen, E.C.M.; Chen, E.S.D. Negative ion mass spectra, electron affinities, gas phase acidities, bond dissociation energies, and negative ion states of cytosine and thymine. *J. Phys. Chem. B* **104**, 7835–7844 (2000)

(10) Sevilla, M.D.; Besler, B.; Colson, A.-O. Ab initio molecular orbital calculations of DNA radical ions. 5. Scaling of calculated electron affinities and ionization potentials to experimental values. *J. Phys. Chem.* **99**, 1060–1063 (1995)

(11) Wetmore, S.D.; Himo, F.; Boyd, R.J.; Eriksson, L.A. Effects of ionizing radiation on crystalline cytosine monohydrate. *J. Phys. Chem. B* **102**, 7484–7491 (1998)

(12) Wetmore, S.D.; Boyd, R.J.; Eriksson, L.A. Radiation products of thymine, 1-methylthymine, and uracil investigated by density functional theory. *J. Phys. Chem. B* **102**, 5369–5377 (1998)

(13) Wetmore, S.D.; Boyd, R.J.; Eriksson, L.A. Comparison of experimental and calculated hyperfine coupling constants. Which radicals are formed in irradiated guanine. *J. Phys. Chem. B* **102**, 9332–9343 (1998)

(14) Oyler, N.A.; Adamowicz, L. Theoretical ab initio calculations of the electron affinity of thymine. *Chem. Phys. Lett.* **219**, 223–227 (1994)

(15) Oyler, N.O.; Adamowicz, L. Electron attachment to uracil. Theoretical ab initio study. *J. Phys. Chem.* **97**, 11122–11123 (1993)

(16) Dolgounitcheva, O.; Zakrzewski, V.G.; Ortiz, J.V. Anionic and neutral complexes of uracil and water. *J. Phys. Chem. A* **103**, 7912–7917 (1999)

(17) Desfrancois, C. *et al* Experimental and theoretical ab initio study of the influence of N-methylation on the dipole-bound electron affinities of thymine and uracil. *J. Chem. Phys.* **110**, 11876–11883 (1999)

(18) Smith, D.M.A.; Jalbout, A.F.; Smets, J.; Adamowicz, L. Cytosine anions: ab initio study. *Chem. Phys.* **260**, 45–51 (2000)

(19) Wetmore, S.D.; Boyd, R.J.; Eriksson, L.A. Electron affinities and ionization potentials of nucleotide bases. *Chem. Phys. Lett.* **322**, 129–135 (2000)

(20) Wesolowski, S.S.; Leininger, M.L.; Pentchew, P.N.; Schaefer, H.F. Electron affinities of the DNA and RNA bases. *J. Am. Chem. Soc.* **123**, 4023–4028 (2001)

(21) Dolgounitcheva, O.; Zakrzewski, V.G.; Ortiz, J.V. Diffuse-bound and valence-bound anions of cytosine. *J. Phys. Chem. A* **105**, 8782–8786 (2001)

(22) Russo, N.; Toscano, M.; Grand, A. Theoretical determination of electron affinity and ionization potential of DNA and RNA bases. *J. Comput. Chem.* **21**, 1243–1250 (2000)

(23) Li, X.; Cai, Z.; Sevilla, M.D. DFT calculations of the electron affinities of nucleic acid bases: Dealing with negative electron affinities. *J. Phys. Chem. A* **106**, 1596–1603 (2002)

(24) Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993)

(25) Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988)

(26) Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996)

(27) Tsuneda, T.; Suzumura, T.; Hirao, K. A new one-parameter progressive Colle–Salvetti-type correlation functional. *J. Chem. Phys.* **110**, 10664–10678 (1999)

(28) Vosko, S.H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. *Can. J. Phys.* **58**, 1200–1211 (1980)

(29) Hertwig, R.H.; Koch, W. On the parameterization of the local correlation functional. What is Becke-3-LYP. *Chem. Phys. Lett.* **268**, 345–351 (1997)

(30) Schmidt, M.W. *et al* General atomic and molecular electronic structure system. *J. Comput. Chem.* **14**, 1347–1363 (1993)

(31) Crawford, O.H. Negative ions of polar molecules. *Mol. Phys.* **20**, 585–591 (1971)

(32) Defrancois, C. *et al* An experimental test for small cluster structure calculations. *J. Chem. Phys.* **102**, 4952–4964 (1995)

(33) Richardson, N.A.; Wesolowski, S.S.; Schaefer, H.F. The adenine-thymine base pair radical anion: Adding an electron results in a major structural change. *J. Phys. Chem. B* **107**, 848–853 (2003)

Chapter 3

Structure and Reactivity Studies of *cis*-Platinum (II) Complexes: Solvent Effect

In this chapter we present our results of DFT calculations on some *cis*-platinum(II) complexes, including clinically used drug molecules, cisplatin, carboplatin, and oxaliplatin and compare the selected geometric parameters with available X-ray data. Global reactivity descriptors viz., chemical hardness, chemical softness, electophilicity index and local reactivity descriptors viz., Fukui function and philicity values of the systems are calculated to understand their reactive nature and reactive sites. Inclusion of solvent effect increases the stability of the complexes. Interestingly, in solvent phase, global descriptors change the trend of reactivity with respect to their trend in the gas phase which correlates well with the available experimental results. A quantitative structure–activity relationship (QSAR) of the complexes has been modelled based on the DFT derived electrophilicity indices against A2780 human ovarian adenocarcinoma cell line to establish the importance of the descriptor in predicting cytotoxicity. [Sarmah and Deka, *Int. J. Quant. Chem.* **108**, 1400–1409 (2008)]

3.1 Introduction

Platinum complexes in cancer chemotherapy deserve a special attention by the landmark discovery of the antitumoural properties of cisplatin.[1] However, as described in Chapter 1, the clinical application of cisplatin is limited by serious side effects, such as nephrotoxicity, neurotoxicity and emetogenesis. Furthermore, cisplatin undergoes many non-selective reactions with a variety of biomolecules, such as proteins and phospholipids.[2] A number of new potential drug candidates were designed in order to overcome these side effects. The cisplatin analogues, carboplatin and oxaliplatin, have been approved for clinical use, and nowadays they are increasingly used in cancer treatment.[3] Correlating antitumour activity, a number of investigations have shown that the most likely target of cisplatin is cellular DNA.[4] Over the last few years, many experimental and theoretical studies have been carried out to provide detailed insight into the binding mechanism of cisplatin-DNA interaction.

The application of DFT, both in its conceptual and computational aspects, to deal a variety of problems in chemistry has increased dramatically in recent years. Global reactivity descriptors introduced within the context of conceptual DFT namely, global hardness (η), global softness (S), electronegativity (χ) and chemical potential (μ), and electrophilicity index (ω) represent the reactivity of a molecule as a whole.[5-9] On the other hand, the local parameters, Fukui function, local softness and philicity (ω^+) describe the relative reactivity and site selectivity in chemical reactions.[10-12] In most of the cases Fukui function values are found to be successful in explaining experimental reactivity sequences of chemical species.[13,14] However, in some cases 'relative electrophilicity', (f_M^+ / f_M^-), and 'relative nucleophilicity', (f_M^- / f_M^+), express the reactivity of atoms in molecules in a better way.[15] The successful applications of these reactivity descriptors have been reported for a variety of chemical compounds and reactions.[16-19] The first application of DFT based reactivity descriptors to a biological system was the inorganic model systems for arsenate reductase and phosphatase.[20] Recently, Beck has calculated Fukui functions for a number of drugs and agrochemicals to investigate whether maxima of these functions can be related to the sites of metabolism.[21] The usefulness of reactivity descriptors, electrophilicity and philicity indices in QSAR modelling to

predict the biological activity/toxicity/property of different systems has been recently reported in a series of papers.[22-24]

Investigation of reactive nature of drug molecules is important to understand their interactions with other molecules in the cellular environment and study of their binding properties. In this chapter, we have calculated global reactivity descriptors to study the reactivity of some model *cis*-platinum (II) complexes including few clinically used platinum anticancer drugs. The list of the studied *cis*-platinum(II) complexes are tabulated in Table 3.1. We have also demonstrated the potential of local reactivity parameters in describing the reactive sites of the molecules. The influence of solvation on the reactivity of the complexes is analyzed. Further, ω values are used to calculate the anticancer activities ($\log(IC_{50}^{-1})$) of the complexes against A2780 human ovarian adenocarcinoma cell line.

Table 3.1: Platinum(II) complexes

Complexes	no.
cis-diamminedicholrideplatinum(II) (cisplatin)	1
cis-diammine(1,1'-cyclobutanedicarboxylato)platinum(II) (carboplatin)	2
cis-(1,1'-cyclobutanedicarboxylato)ethylenediamineplatinum(II)	3
cis-[(1R,2R)-1,2-diaminocyclohexane]oxalatoplatinum(II) (oxaliplatin)	4
cis-[(1R,2R)-1,2-diaminocyclohexane]malonatoplatinum(II)	5
cis-dichloride[(1R,2R)-1,2-diaminocyclohexane]platinum(II)	6
cis-(1,1'-cyclobutanedicarboxylato)[(1R,2R)-1,2-diaminocyclohexane]platinum(II)	7

3.2 Methodology

All the DFT calculations were carried out using the DMol3 program.[25] DNP basis set in combination with three generalized gradient approximation (GGA) exchange-correlation functionals BLYP, BOP and HCTH were used for geometry optimizations. The DNP basis set is of double numeric quality augmented with polarization functions. DMol3 has a facility that applies scalar relativistic corrections to this basis set. We performed all electron calculations, including relativistic effects for all complexes. The most popular exchange-correlation BOP is

functional BLYP is the combination of the exchange functional developed by Becke[26] with the gradient corrected correlation functional of Lee, Yang and Parr.[27] BOP is the combination of Becke[26] exchange and one-parameter progressive OP[28] correlation functionals. In HCTH both exchange and correlation parts were developed by Hamprecht, Cohen, Tozer and Handy.[29] The minimized structures were verified as local minima on the potential energy surface by normal-mode frequency analysis. The global reactivity descriptors chemical potential, hardness and electrophilicity index were calculated for all the systems. Hirshfeld charges[30] were used for computing the condensed Fukui functions. While comparing the calculated geometrical parameters of the drug molecules with the experimental data, we found that the BLYP/DNP calculations worked better than BOP/DNP and HCTH/DNP calculations. Hence, we used BLYP/DNP level calculations to study the effect of solvent (water) using COSMO[31] with dielectric constant of 78.39. Liner regression analysis (LRA) was performed[32] using least square fitting to derive QSAR model of $\log(IC_{50}^{-1})$ using the calculated electrophilicity indices.

3.3 Results and Discussion

3.3.1 Optimized Structure

The gas phase optimized structures obtained from BLYP/DNP level calculations for the platinum (II) complexes are shown in Figure 3.1. Selected bond lengths and bond angles of the complexes, both in gas and solvent phases are provided in Table 3.2. The optimized geometry of cisplatin (1) has a square-planar structure in both gas and solvent media, which is in agreement with X-ray crystal structure reported by Milburn and Truter.[33] The calculated Pt–Cl and Pt–N bond lengths in gas phase are 2.32 Å and 2.11Å, respectively in accordance with their experimental values (Table 3.2). In the gas phase calculations the N—Pt—N angle (97.1°) deviates slightly more from its experimental value (87°±1.5°). The Cl—Pt—Cl angle is also larger than its experimental value by about 5–6.° Similar deviation of bond angles in cisplatin was reported in theoretical studies performed by Wysokinski and Michalska.[34] These higher values could be due to the intramolecular N—H...Cl interaction of cisplatin in the gas phase, which opens up both the N—Pt—N and Cl—Pt—Cl angles. Inclusion of solvent effect decreases the N—Pt—N and Cl—Pt—Cl angles from their gas phase values to 89.2° and 93.9°, respectively.

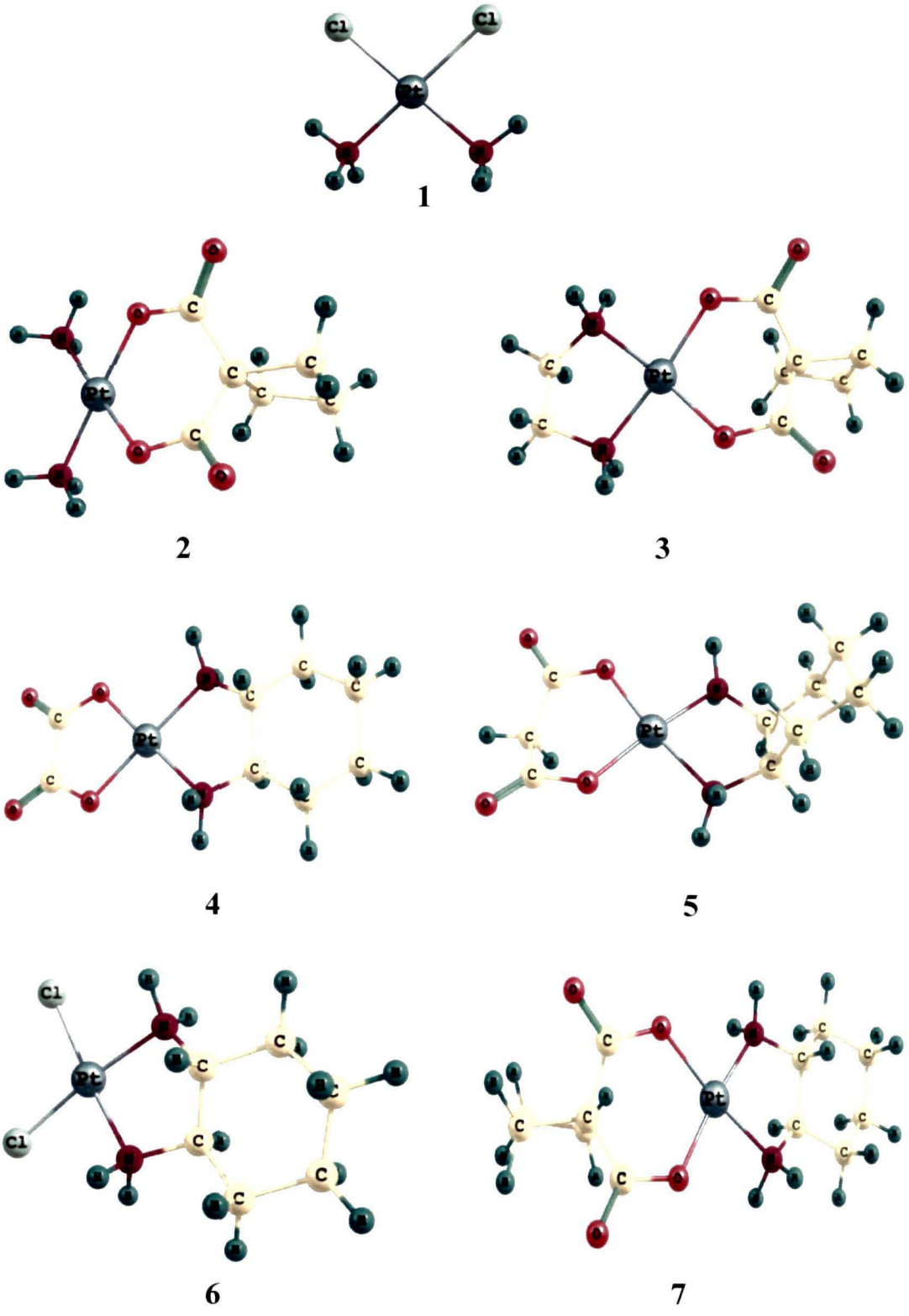


Figure 3.1: Optimized structures for the *cis*-Pt(II) complexes obtained from BLYP/DNP calculations.

In carboplatin (2) chlorine atoms are replaced by cyclobutanedicarboxylate groups where two ammine groups are in the *cis* position. The planar environment of the platinum atom and the boat conformation for the six-membered chelate ring obtained from gas phase and solvent phase calculations are in agreement with the X-ray diffraction data[35] (Table 3.2). Our calculations both in gas and solvent media predict a slightly puckered structure of cyclobutane ring which is also supported by the experimental data. The Pt—N bond length is slightly shorter than that of cisplatin. This is in consistent with the results obtained from different theoretical methods by Wysokinski and Michalska.[34] The geometrical parameters of the complex 3, in which carrier ligands NH_3 of carboplatin are replaced by ethylenediamine, are compared with that of carboplatin in Table 3.2. For this complex, again we have obtained boat conformation of the six-membered chelate ring with a puckered structure of cyclobutane ring.

Oxaliplatin (4) has (1R,2R)-1,2-diaminocyclohexane (DACH) as a carrier ligand and oxalate (OX) as a leaving ligand. From complex 5–7, keeping DACH as a carrier ligand, the leaving ligands are varied as malonato (MAL), chloride and 1,1'-cyclobutanedicarboxylato (CBDC), respectively. The bond lengths and bond angles of oxaliplatin obtained from gas and solvent phases are close to its X-ray crystal structure data.[36] In both the phases, chair configuration of the cyclohexane ring with two amino groups in equatorial positions is found in accordance with the experimental results. The geometry of complex 5 is compared with the experimental result[36] in Table 3.2. In this compound, the cyclohexane ring has chair conformation whereas in direct contrast, the malonato ligand shows a boat conformation for the six membered Pt—O—C—C—C—O ring. This is in good agreement with the experimental data. For compounds 6 and 7, again we have found cyclohexane ring with chair conformation whereas the carboxylato group of CBDC in 7 has boat conformation. Their geometrical parameters are comparable with the X-ray data of oxaliplatin.[36] Thus, it is observed that our calculated geometries for all the complexes are in good agreement with their experimental data. However slightly higher values of bond lengths for all most all the cases are due to the BLYP exchange correlation functional that is known to overestimate the bond lengths.[37,38]

Table 3.2: Calculated bond lengths (in Å) and angles (in deg) for all complexes at BLYP/DNP level. Solvent phase values are in parenthesis.

	X-ray (1)	1		X-ray (2)	2	3		X-ray (4)	X-ray (5)	4	5	6	7
Pt—Cl	2.33±0.01	2.32	Pt—N1	2.01	2.10	2.09	Pt—O1	2.01	2.02	2.00	2.01		2.00
		(2.33)			(2.07)	(2.07)				(2.04)	(2.03)		(2.04)
Cl—Pt—Cl	91.9±0.3	96.8	Pt—O1	2.02	2.00	2.00	Pt—N1	2.06	2.03	2.08	2.07	2.07	2.07
		(93.9)			(2.03)	(2.04)				(2.06)	(2.06)	(2.07)	(2.06)
Pt—N	2.01±.04	2.11	O1—Pt—N1	88.2	80.6	90.3	O1—C7	1.21	1.24	1.27	1.31		1.32
		(2.08)			(88.4)	(82.9)				(1.25)	(1.29)		(1.29)
N—Pt—N	87±1.5	97.1	N1—Pt—N2	93.6	103.1	84.1	O2—C8	1.32		1.35			
		(89.2)			(91.4)	(93.7)				(1.32)			
			O1—Pt—N2	177.9	175.8	173.9	O2—C9		1.35		1.34		1.34
					(179.9)	(176.2)					(1.32)		(1.32)
			O1—Pt—O2	88.9	95.9	95.4	Pt—Cl1					2.32	
					(91.5)	(89.7)						(2.35)	
							O2—Pt—O1	82.5	90.3	84.3	96.0		94.9
										(82.3)	(91.8)		(89.2)
							N1—Pt—O1	96.0	93.2	96.4	89.9		90.7
										(98.6)	(92.8)		(94.5)
							N2—Pt—O1	175.6	177.0	179.1	173.6		173.8
										(178.7)	(174.7)		(176.4)
							N1—Pt—O2	169.7	176.4	172.8	173.3		174.2
										(179.7)	(175.0)		(176.1)

3.3.2 *Reactivity*

3.3.2.1 *Global descriptors*

The gas phase values of chemical hardness (η), chemical potential (μ) and electrophilicity index (ω) computed using DNP basis set in connection with the three exchange-correlation functionals for the compounds are given in Table 3.3. It is observed that all the methods predict similar trend for these descriptors except the electrophilicity trend for oxaliplatin and complex 5 at BLYP/DNP level. According to the maximum hardness principle (MHP)[39,40], the most stable structure has the maximum hardness. Thus in gas phase, complex 7 is found to be the most stable structure. This complex has the minimum value of ω and maximum value of μ and hence the least reactive whereas cisplatin (1) has the maximum ω and minimum μ values and is the most reactive among the complexes with a minimum value of η. The higher reactivity of cisplatin agrees well with most of the platinum drug literatures, for which the toxicity of the molecule is more.

Table 3.3: Calculated hardness (η, in eV), chemical potential (μ, in eV), and electrophilicity index (ω, in eV) for all complexes in gas phase.

	BLYP			BOP			HCTH		
Complex	η	μ	ω	η	μ	ω	η	μ	ω
1	1.474	-3.491	4.134	1.485	-3.371	3.826	1.605	-3.627	4.098
2	1.714	-3.398	3.368	1.703	-3.263	3.126	1.813	-3.539	3.454
3	1.690	-3.165	2.964	1.682	-3.036	2.740	1.792	-3.336	3.105
4	1.629	-2.995	2.753	1.626	-2.870	2.533	1.776	-3.182	2.850
5	1.676	-3.034	2.746	1.683	-2.925	2.542	1.798	-3.209	2.864
6	1.504	-3.040	3.072	1.506	-2.931	2.852	1.626	-3.212	3.172
7	1.724	-2.999	2.608	1.716	-2.871	2.402	1.824	-3.158	2.734

The variation of hardness calculated with BLYP/DNP level for the complexes is shown in Figure 3.2. The leaving ligands of each complex are also displayed in the

plot. It is observed that the compounds having chloride groups (1,6) as leaving ligand exhibit higher reactivity whereas the CBDC as leaving groups make the complexes harder.

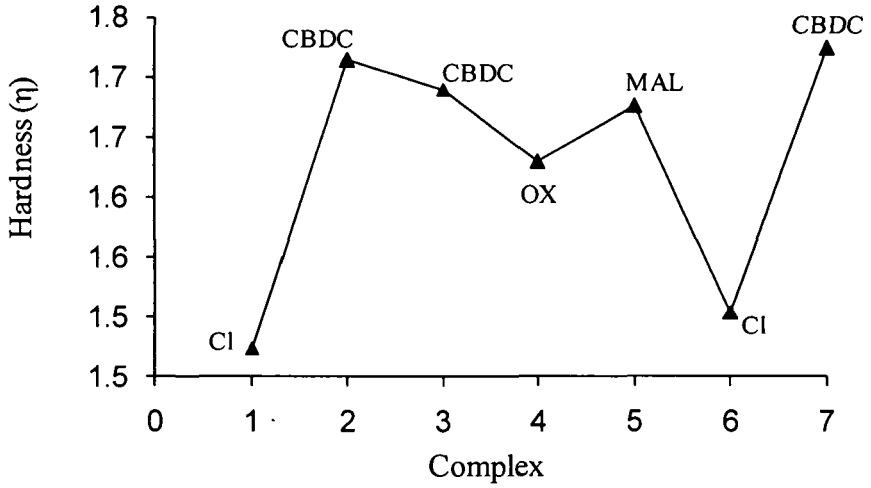


Figure 3.2: Variation of hardness of the molecules calculated at BLYP/DNP level. The leaving ligands of the complexes are also shown.

The calculated values of global reactivity descriptors, η, μ and ω from the BLYP/DNP method in solvent medium are presented in Table 3.4. The reactivity trend of the molecules in gas phase changes with the inclusion of solvent medium. The variation of chemical hardness values of the complexes in both the media is shown in Figure 3.3. The higher values of hardness in solvent-phase indicate that all complexes display increased stabilization in this medium. However, only slight difference in hardness is observed for oxaliplatin (4). It is very interesting to note from Figure 3.3 that the most reactive complex changes from cisplatin (1) to oxaliplatin (4) in going from gas phase to solvent phase. On the other hand, solvent-phase calculations predict carboplatin (2) as the least reactive among the other platinum complexes in contrast to the gas phase calculations. The higher reactivity of oxaliplatin than that of other two clinically used drugs (2 and 4) is in agreement with the data reported by Screnci *et al.*[41] They observed reactivity of some platinum anticancer drugs by measuring the half-life of binding to plasma proteins in vitro and found higher reactivity of oxaliplatin than that of cisplatin.

Table 3.4: Calculated hardness (η, in eV), chemical potential (μ, in eV), and electrophilicity index (ω, in eV) for all complexes in solvent phase at BLYP/DNP level.

complex	BLYP/DNP		
	η	μ	ω
1	1.685	-3.618	3.884
2	1.973	-3.336	2.851
3	1.803	-3.101	2.666
4	1.638	-3.777	4.357
5	1.957	-3.407	2.966
6	1.668	-3.573	3.815
7	1.950	-3.536	3.748

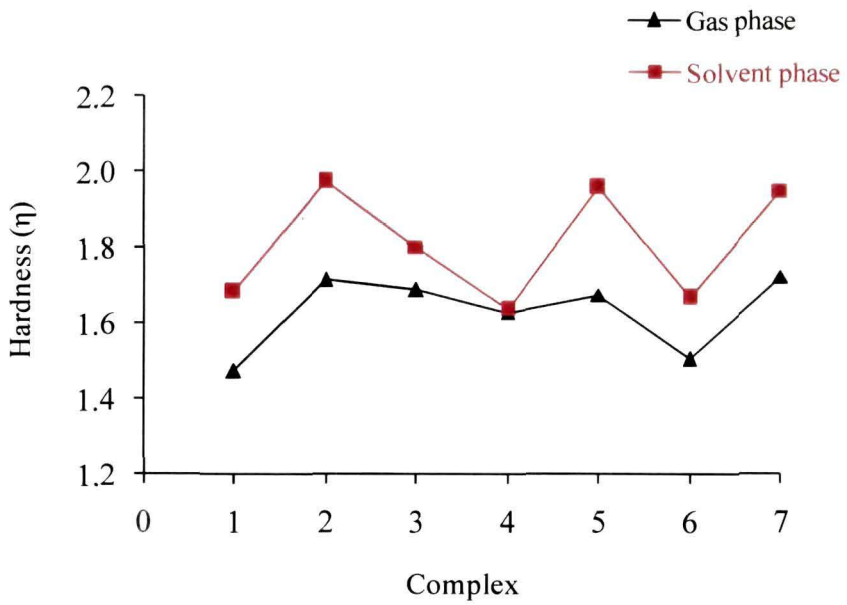


Figure 3.3: Variation of hardness of the molecules for both gas and solvent media.

Figure 3.4 shows the variation of electrophilicity descriptors (ω) for all selected systems both in gas and solvent media. It is interesting to observe that complexes 1–3 have lower values of ω in the solvent medium compared to that in the gas phase; however, this trend becomes reverse for complexes 4–7. This variation indicates that the presence of DACH as carrier groups (4–7) exhibited the highest electrophilicity values, which could be a reason behind higher cytotoxic effects of these complexes.

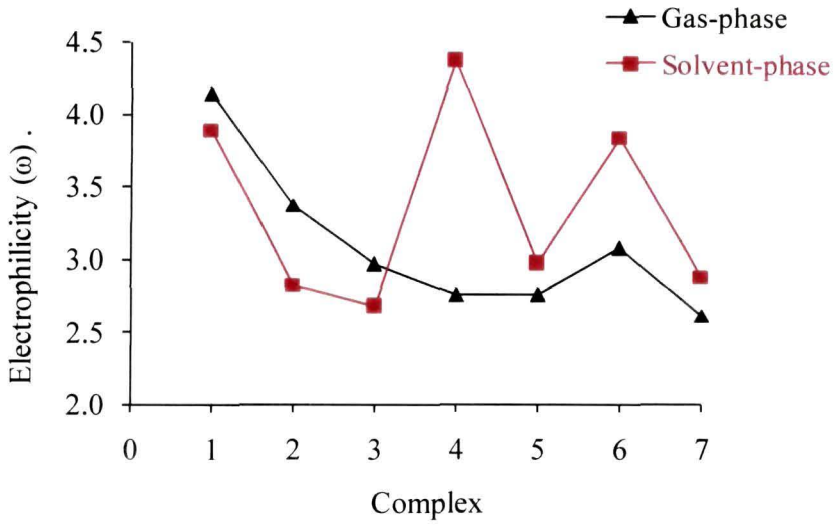


Figure 3.4: Variation of electrophilicity index of the molecules for both gas and solvent media.

3.3.2.2 Local descriptors

Local reactivity parameters describe the relative reactivity and site selectivity of chemical species. The nucleophilic attack at a particular site of a system represents the site with maximum value of Fukui function, f^+ and/or local philicity, ω^+. Similarly, electrophilic attack at a particular site of a system represents the site with maximum value of Fukui function, f^- and/or local philicity, ω^-. The platinum anticancer drug-DNA binding mechanism involves the nucleophilic attack at Pt atom. So, we have calculated f^+ and ω^+ values for all the atoms of the complexes both in gas and solvent media. As expected, in all the systems Pt sites are prone to nucleophilic attack with maximum values of f^+ and ω^+. The gas phase and solvent

phase values of f^+ and ω^+ of Pt atom in complexes (1–7) are presented in the Table 3.5.

Table 3.5: Fukui function (f^+) and local philicity indices (ω^+) of Pt atom of the complexes both in solvent and gas phases calculated at BLYP/DNP level.

	f^+		ω^+	
system	Solvent phase	Gas phase	Solvent phase	Gas phase
1	0.272	0.210	1.056	0.868
2	0.245	0.162	0.691	0.546
3	0.242	0.145	0.645	0.430
4	0.089	0.126	0.388	0.347
5	0.274	0.150	0.813	0.412
6	0.260	0.196	0.992	0.602
7	0.236	0.132	0.885	0.344

It is seen from Table 3.5 that f^+ and ω^+ values of the Pt atoms of all most all the complexes in the solvent medium are larger than their gas phase values. Solvent phase derived f^+ value of oxaliplatin (4) does not follow the regular trend. However, for this complex ω^+ value in the solvent phase is higher than that of gas phase. These results indicate that the Pt atoms of the complexes in solvent phase are chemically softer than that of gas phase.

3.3.3 Structure-Activity Analysis

The inhibitory activity (IC_{50}) values of the platinum complexes against A2780 human ovarian adenocarcinoma cell line was taken from the results reported by Monti *et al.*[42] As usual, these data were changed to $\log(IC_{50}^{-1})$ to be practical use in the QSAR analysis which is modeled by using a linear regression technique. Linear regression analyses were carried out for both solvent and gas phases by considering the experimental cytotoxic activity $\log(IC_{50}^{-1})$ as a dependent variable

and the DFT based global reactivity descriptors (ω) obtained from BLYP/DNP method as an independent variable. It is observed that in the gas phase ω fails to predict the cytotoxic activity of the complexes. In direct contrast, the ω values in the solvent phase could predict the cytotoxicity of the complexes. The modeled regression equation in the solvent phase is given by

$$\log(IC_{50}^{-1}) = -6.152 + (1.565 \pm 0.74)\,\omega \qquad (3.1)$$
$$N=7,\ r^2 = 0.889,\ SD = 0.389$$

It is found that the derived model accounts for the 88.9% variance in the experimental data with a root-mean-square error (SD) of 0.389. Table 3.6 lists the experimental and calculated cytotoxicity of the complexes.

Table 3.6: Experimental and solvent phase calculated cytotoxic activity ($\log(IC_{50}^{-1})$) values of the *cis*-Pt(II) complexes against A2780 human ovarian adenocarcinoma cell line.

	$\log(IC_{50}^{-1})$		
system	observed[a]	calculated	residual[b]
1	-0.315	-0.070	-0.245
2	-2.036	-1.737	-0.300
3	-0.030	-0.178	0.149
4	-2.220	-1.977	-0.243
5	-0.798	-1.509	0.711
6	0.734	0.671	0.063
7	-0.419	-0.284	-0.135

a) Ref. 42

b) Difference between the experimental and calculated values of log (IC_{50}^{-1})

As observed in the plot of Figure 3.5, the calculated cytotoxicity values, computed from the equation above, are close to the experimental values. This finding indicates that the electrophilicity indices (ω) are capable of predicting the cytotoxicity in a reasonable way with a correlation coefficient (r) of 0.943. Thus, $\log(IC_{50}^{-1})$ exhibits

a reasonable correlation with electrophilicity which clearly indicates the fact that electrophilicity can be successfully used as descriptors in the prediction of cytotoxicity of platinum complexes.

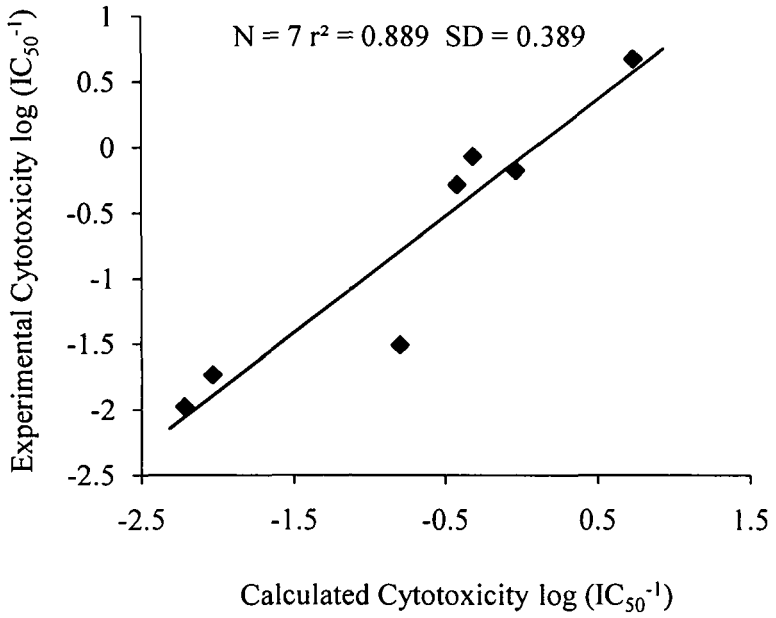


Figure 3.5: A plot between experimental and calculated log(IC_{50}^{-1}) values of all *cis*-Pt(II) complexes against A2780 human ovarian adenocarcinoma cell line

3.4 Conclusions

In this study, we have carried out systematic theoretical calculations of *cis*-platinum (II) complexes both in gas and solvent phases in order to assess their structure, stability and reactivity. The geometrical parameters of the molecules agree well with the available experimental data. The reactivity sequences derived by calculating global hardness values are different in the gas phase and solvent medium. Global hardness values calculated in the solvent medium reproduced the experimental reactivity trend of cisplatin, carboplatin and oxaliplatin. QSAR analysis performed using solvent phase derived electrophilicity values showed a good correlation with the experimental cytotoxicity values. Thus, these results indicate the solvent dependence reactivity natures of platinum drugs as well as significance of electrophilicity index in predicting structure-activity relationship of the molecules.

References

(1) Rosenberg, B.; Van Camp, L.; Krigas, T. Inhibition of cell division in Escherichia Coli by electrolysis products from a platinum electrode. *Nature* **205**, 698–699 (1965)

(2) Wang, K.; Lu, J.F.; Li, R.C. The events that occur when cisplatin encounters cells. *Coord. Chem. Rev.* **151**, 53–88 (1996)

(3) Boulikas, T.; Vougiouka, M. Cisplatin and platinum drugs at the molecular level. *Oncol. Rep.* **10**, 1663–1682 (2003)

(4) Lippard, S.J. New chemistry of an old molecule: cis-$[Pt(NH_3)_2Cl_2]$ *Science*, **218**, 1075–1082 (1982)

(5) Pearson, R.G. Hard and soft acids and bases. *J. Am. Chem. Soc.*, **85**, 3533–3539 (1963)

(6) Pearson, R.G. Absolute electronegativity and absolute hardness of Lewis acids and bases. *J. Am. Chem. Soc.* **107**, 6801–6806 (1985)

(7) Parr, R.G.; Donnelly, R.A.; Levy, M.; Palke, W.E. Electronegativity: The density functional viewpoint. *J. Chem. Phys.* **68**, 3801–3807 (1978)

(8) Parr, R.G.; Pearson, R.G. Absolute hardness: Companion parameter to absolute electronegativity. *J. Am. Chem. Soc.* **105**, 7512–7516 (1983)

(9) Parr, R.G.; Szentpaly, L.V.; Liu, S. Electrophilicity Index. *J. Am. Chem. Soc.* **121**, 1922–1924 (1999)

(10) Parr, R.G.; Yang, W. Density functional approach to the frontier-electron theory of chemical reactivity. *J. Am. Chem. Soc.* **106**, 4049–4050 (1984)

(11) Yang, W.; Parr, R.G. Hardness, softness, and the fukui function in the electronic theory of metals and catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6723–6726 (1985)

(12) Chattaraj, P.K.; Maiti, B.; Sarkar, U. Philicity: A Unified treatment of chemical reactivity and selectivity. *J. Phys. Chem. A* **107**, 4973–4975 (2003)

(13) Langenaeker, W.; De Decker, M.; Geerlings, P.; Raeymaekers, P. Quantum-chemical study of the Fukui function as a reactivity index: probing the acidity of bridging hydroxyls in zeolite-type model systems. *J. Mol. Struct. (THEOCHEM).* **207**, 115–130 (1990)

(14) Ayers, P.W.; Parr, R.G. Variational principles for describing chemical reactions: The Fukui function and chemical hardness revisited. *J. Am. Chem. Soc.* **122**, 2010–2018 (2000)

(15) Roy, R.K.; Krishnamurti, S.; Geerlings, P.; Pal, S. Local softness and hardness based reactivity descriptors for predicting intra and intermolecular reactivity sequences: Carbonyl compounds. *J. Phys. Chem. A* **102**, 3746–3752 (1998)

(16) Chandrakumar, K.R.S.; Pal, S. Study of local hard-soft acid-base principle to multiple-site interactions. *J. Phys. Chem. A* **106**, 5737–5744 (2002)

(17) Chandrakumar, K.R.S.; Pal, S. DFT and local reactivity descriptor studies on the nitrogen sorption selectivity from air by sodium and calcium exchanged Zeolite-A", *Colloids and Surfaces A* **205**, 127–138 (2002)

(18) Chandrakumar, K.R.S.; Pal, S. The concept of density functional theory based descriptors and its relation with the reactivity of molecular systems: A semi-quantitative study. *Int. J. Mol. Sci.* **3**, 324–337 (2002)

(19) Pal, S.; Chandrakumar, K.R.S. Critical study of local reactivity descriptors for weak interactions: Qualitative and quantitative analysis of adsorption of molecules in the Zeolite lattice. *J. Am. Chem. Soc.* **122**, 4145–4153 (2000)

(20) Roos, G.; Loverix, S.; De Proft, F.; Wyns, L.; Geerlings, P. A Computational and conceptual DFT study of the reactivity of anionic compounds: Implications for enzymatic catalysis. *J. Phys. Chem. A* **107**, 6828–6836 (2003)

(21) Beck, M.E. Do Fukui function maxima relate to sites of metabolism? A critical case study. *J. Chem. Inf. Model.* **45**, 273–282 (2005)

(22) Padmanabhan, J.; Parthasarathi, R.; Subramanian, V.; Chattaraj, P.K. Group philicity and electrophilicity as possible descriptors for modeling ecotoxicity applied to chlorophenols. *Chem. Res. Toxicol.* **19**, 356–364 (2006)

(23) Roy, D.R.; Parthasarathi, R.; Maiti, B.; Subramanian, V.; Chattaraj, P.K. Electrophilicity as a possible descriptor for toxicity prediction. *Bioorg. Med. Chem.* **13**, 3405–3412 (2005)

(24) Parthasarathi, R.; Subramanian, V.; Roy. D.R.; Chattaraj, P.K. Electrophilicity index as a possible descriptor of biological activity. *Bioorg. Med. Chem.* **12**, 5533–5543 (2004)

(25) Delley, B. An all-electron numerical method for solving the local density functional for polyatomic molecules. *J. Chem. Phys.* **92**, 508–517 (1990)

(26) Becke, A.D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988)

(27) Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev.* B **37**, 785–789 (1988)

(28) Tsuneda, T.; Suzumura, T.; Hirao, K. A new one-parameter progressive Colle–Salvetti-type correlation functional. *J. Chem. Phys.* **110**, 10664–10678 (1999)

(29) Hamprecht, F.A.; Cohen, A.J.; Tozer, D.J.; Handy, N.C. Development and assessment of new exchange–correlation functional. *J. Chem. Phys.* **109**, 6264–6271 (1998)

(30) Hirshfeld, F.L. The theoretical tool used is the minimum entropy deficiency principle (minimum missing information principle. *Theor. Chim. Acta.* **44**, 129–138 (1977)

(31) Andzelm, J.; Koelmel, C.; Klamt, A. Incorporation of solvent effects into density functional calculations of molecular energies and geometries. *J. Chem. Phys.* **103**, 9312–9320 (1995)

(32) MATLAB; The Math Works, Inc.: Natick, USA. 1999.

(33) Milburn, G.H.W.l Truter, M.R. The crystal structures of cis- and trans-dichlorodiammineplatinum (II). *J. Chem. Soc. A* 1609–1616 (1966)

(34) Wysokiński, R.; Michalska, D. The Performance of different density functional methods in the calculation of molecular structures and vibrational spectra of platinum(II) antitumor drugs: Cisplatin and carboplatin. *J. Comput. Chem.* **9**, 901–912 (2001)

(35) Beagley, B. *et al* The crystal and molecular structure of *cis*-diammine-1,1-cyclobutanedicarboxoplatinum(II) [*cis*-$Pt(NH_3)_2CBDCA$]. Dynamic puckering of the cyclobutane ring. *J. Mol. Struct.* **130**, 97–102 (1985)

(36) Bruck, M.A.; Bau, R.; Noji, M.; Inagaki, K.; Kidani, Y. The Crystal structures and absolute configurations of the anti-Tumor complexes Pt(oxalato)(1R,2R-cyclohexanediamine) and Pt(malonato)(1R,2R-cyclohexanediamine). *Inorg. Chim. Acta.* **92**, 279–284 (1984)

(37) Spiegel, K.; Rothlisberger, U.; Carloni, P. Cisplatin binding to DNA oligomers from hybrid Car-Parrinello/molecular dynamics simulations. *J. Phys. Chem. B* **108**, 2699–2707 (2004)

(38) Carloni, P. *et al* Structure and bonding in cisplatin and other Pt(II) complexes. *Chem. Phys. Lett.* **234**, 50–56 (1995)

(39) Parr, R.G.; Chattaraj, P.K. Principle of maximum hardness. *J. Am. Chem. Soc.* **113**, 1854–1855 (1991)

(40) Chattaraj, P.K.; Liu, G.H.; Parr, R.G. The maximum hardness principle in the Gyftopoulos-Hatsopoulos three-level model for an atomic or molecular species and its positive and negative ions. *Chem. Phys. Lett.* **237**, 171–176 (1995)

(41) Screnci, D. *et al* Relationships between hydrophobicity, reactivity, accumulation and peripheral nerve toxicity of a series of platinum drugs. *British. J. Cancer.* **82**, 966–972 (2000)

(42) Monti, E. *et al* Cytotoxicity of *cis*-platinum(II) conjugate models. The effect of chelating arms and leaving groups on cytotoxicity: A quantitative structure-activity relationship approach. *J. Med. Chem.* **48**, 857–866 (2005)

Chapter 4

QSAR and QSPR Studies of Several *cis*-Platinum Complexes

In the present chapter, quantitative structure-activity relationship (QSAR) analyses of *cis*-platinum complexes with different leaving and carrier ligands against parental and resistant ovarian cancer cell lines have been performed using conceptual density functional theory (DFT). The calculated QSAR models become more significant with incorporation of solvent effects indicating its importance in studying biological activity. Given the importance of logarithimic n-octanol/water partition coefficient ($\log P_{o/w}$) in drug metabolism and cellular uptake, we modelled the $\log P_{o/w}$ of 24 *cis*-platinum complexes by the quantitative structure-property relationship (QSPR) analyses against five different concentrations of MeOH using DFT and molecular mechanics (MM) derived descriptors. Using the same descriptors, QSPR analysis has been performed for an additional set of 20 platinum complexes and the predicted ability of the model is investigated by calculating $\log P_{o/w}$ of 4 compounds in the test set. The solvent medium played an important role in QSPR analysis by increasing the internal predictability of the models. [Sarmah and Deka, *J. Compt. Aided. Mol. Des.* **23**, 343–354 (2009)]

4.1 Introduction

The inorganic drug cisplatin is widely used in the treatment of a number of solid malignancies through triggering apoptotic pathways in the tumour cells.[1] It is used for treatment of testicular, ovarian, head and neck, bladder, and small-cell lung cancer.[2,3] Despite its success, cisplatin has several disadvantages that include nephrotoxicity, neurotoxicity, and ototoxicity. Furthermore, development of resistance in some tumors such as colorectal and nonsmall cell lung cancers, after continued drug administrations presents a serious complication in chemotherapy. The limitations of usefulness of cisplatin have stimulated research toward developing analogues of cisplatin with lesser toxic effects. The cisplatin analogues, carboplatin and oxaliplatin, have been approved for clinical use, and nowadays they are increasingly used in cancer treatment. In continued search of new platinum-based drugs of improved anticancer activity, more than 3000 platinum compounds have been prepared and tested against several tumor cell lines. The cytotoxicity of platinum complexes depends on the nature of carrier and leaving ligands. Monti *et al.*[4] studied the cytotoxicities of 16 platinum complexes with different leaving and carrier groups in two cancer cell lines. Their results confirm the Cleare and Hoeschele's empirical rules that the presence of NH_3 and DACH [(1R,2R)-1,2-diaminocyclohexane] as carrier groups and chloride and oxalate as leaving ligands yield the highest cytotoxic effects.

Hydrophobicity, measured as logarithm of the 1-octanol/water partition coefficient ($\log P_{o/w}$) is a very important property owing to its usefulness to assess biological effects relevant to drug action, such as lipid solubility, tissue distribution, receptor binding, cellular uptake, metabolism, and bioavailability. This parameter is a useful tool in the field of quantitative structure–activity relationships (QSARs) for several biological effects. Numerous experimental and theoretical studies have been devoted in determining hydrophobicity of different organic molecules of pharmacological and toxicological interest. However, limited studies have been carried out on $\log P_{o/w}$ of platinum complexes. Screnci *et al.*[5] reported $\log P_{o/w}$ of 8 platinum drugs, including 4 platinum (IV) drug molecules using shake-flask method. They also derived another hydrophobicity parameter, $\log k_w$ and observed a weak correlation between $\log P_{o/w}$ and $\log k_w$. Platts *et al.*[6] calculated hydrophobicity of a

series of 24 platinum complexes with the help of RP-HPLC technique and found a good correlation of these values with that derived from DFT calculations.

Although several theoretical calculations have been performed to understand the structure and binding mechanism of platinum drugs with DNA, very few studies have paid attention on QSAR/QSPR analyses of these molecules.[4,6,7,8] Platts *et al.*[7] showed that surface areas of polar and nonpolar atoms of a set of 23 platinum drugs can accurately model the hydrophobicity of the drugs. They also found a good correlation of their calculated hydrophobicity of 5 drugs with the cell uptake. Costa *et al.*[8] calculated rate constants (k) for the first step hydrolysis of a set of platinum(II) complexes containing different electron donating and withdrawing groups and found that values of $\log(k_X/k_H)$ correlated with the electronic Hammett constant σ_m and σ_p. For the first time they have used the values of $\log(k_X/k_H)$ to construct QSAR model of platinum complexes and concluded that the analogues containing substituents which lead to faster hydrolysis can be less mutagenic.

From the previous chapter we found that solvent phase derived elctrophilicity values of platinum complexes can predict their anticancer activity accurately. In continuation, in the present chapter, we have performed QSAR analyses of a set of *cis*-platinum complexes in both gas and solvent phases. In addition to correlate activity, we have intended to calculate hydrophobicity of the complexes from QSPR analysis.

4.2 Computational Details

Full unconstrained geometry optimizations of all complexes were carried out at gradient corrected density functional theory with the program DMol^3 using DNP basis set and BLYP functional. We performed all electron calculations, including relativistic effects for all complexes, as available in DMol^3. All complexes were characterized as minima (no imaginary frequency) in their potential energy surface through harmonic frequency analysis. The reactivity descriptors chemical potential (μ), hardness (η), electrophilicity index (ω), and local philicity (ω^+) were calculated for all the systems. The Hirshfeld[9] population analysis (HPA) was used to calculate the Fukui function. The Conductor-like Screening Model (COSMO) with dielectric constant of 78.4 was adopted to study the solvent (water) effect. The molar refractivity parameter of carrier ligands and surface area of each complex were

obtained from the MM+ computations with HyperChem software.[10] The predictive ability of models was determined using the "leave one out" (LOO) cross-validation method.

4.3 QSAR/QSPR Modelling

From the results of DFT calculations, different descriptors were selected for QSAR and QSPR modelling such as, the energy of highest occupied molecular orbital (E_{HOMO}), energy of lowest unoccupied molecular orbital (E_{LUMO}), energy of the next lowest unoccupied molecular orbital (E_{NL}), energy difference between LUMO and HOMO ($\Delta_{L\text{-}H}$), dipole moments, electrophilicity (ω), hardness (η), philicity (ω^+) etc. In addition, the molecular mechanics parameters such as molar refractivity of carrier ligand (MR_{CL}), van der Waals surface area (SA), molecular volume, hydrophobicity of carrier ligand ($logP_{CL}$) were also selected.

The anticancer activity data of 16 platinum complexes against the A2780 human ovarian adenocarcinoma cell line and its cisplatin resistant subline (A2780Cp8) were taken from the results reported by Monti *et al.*[4] These values were conventionally transformed to $\log(IC_{50}^{-1})$ in QSAR studies. The analyses were performed in both gas and solvent media for the 16 platinum complexes. We carried out QSPR studies to analyze the $\log P_{o/w}$ values of 24 platinum complexes for 0% (extrapolated), 20%, 30%, 40%, and 50% MeOH.[6] Since the partition behavior markedly depends on the solvent, we also performed multiple regression analysis using solvent phase predicted molecular properties. The $\log P_{o/w}$ values of a training set of 20 platinum complexes were also modelled with the same descriptors. Further we investigated predictability of the models by calculating $\log P_{o/w}$ of 4 compounds in the test set. The regression analyses were also performed on 24 complexes obtained from combination of training set and test set. The descriptors having greater correlation to $\log(IC_{50}^{-1})$ and $\log P_{o/w}$ with smaller autocorrelation were selected out to perform the stepwise multiple linear regression. Three parameter QSAR and four parameter QSPR[11] were performed using least square error estimation method[12] to calculate and compare the cytotoxicity ($\log(IC_{50}^{-1})$) and hydrophobicity ($\log P_{o/w}$) of the complexes, respectively. The predictive ability of models was determined using the "leave one out" (LOO) cross-validation method.

4.4 Results and discussion

The platinum complexes used in QSAR analysis are presented in Figure. 4.1.

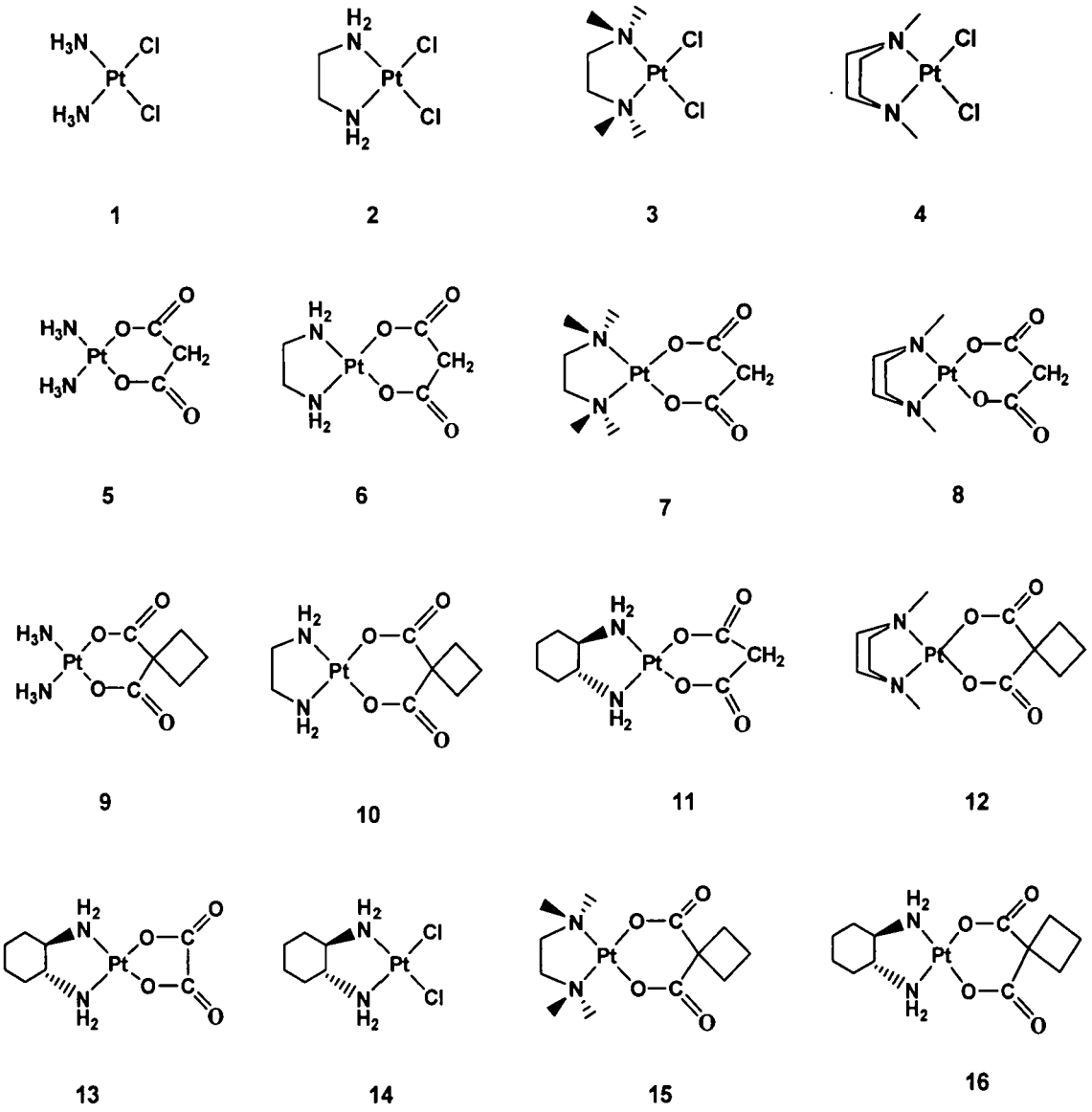


Figure 4.1: Sketch of the platinum complexes used to build QSAR models.

The optimized geometries of all complexes have square planar configuration with angles close to the ideal values of 90° and 180°. The optimized geometry of cisplatin (1) is in good agreement with X-ray crystal structure reported by Milburn and Truter.[13] The calculated Pt–Cl and Pt–N bond lengths are 2.32 Å and 2.11Å, respectively in accordance with their experimental values. The N—Pt—N angle (97.1°) and Cl—Pt—Cl (96.8°) angle are larger by about 5–7° from their

experimental values 87°±1.5° and 91.9°±0.4°, respectively. The planar environment of the platinum atom and the boat conformation for the six-membered chelate ring obtained for carboplatin (9) are in agreement with the X-ray diffraction data.[14] The Pt–N (2.10 Å) and Pt–O (2.00 Å) bond lengths and N—Pt—N angle (96.4°) and O—Pt—O angle (84.3°) of oxaliplatin (13) are close to its X-ray crystal structure data.[15] Chair configuration of the cyclohexane ring with two amino groups in equatorial positions is found in accordance with the experimental results. The geometrical parameters of other complexes, for which X-ray data are not available, are therefore, compared with geometries of the above mentioned drug molecules. We found that our calculated geometries for all complexes are in good agreement with the available experimental data.

4.4.1 QSAR Analysis on A2780 Cell Line

The QSAR equations with absolute values of statistical parameters in both gas and solvent phases for 16 platinum complexes against A2780 cell line are represented by equations 4.1 and 4.2. The values are calculated by considering the $\log(IC_{50}^{-1})$ as a dependent variable and electrophilicity (ω), philicity (ω^+), and energy of the next LUMO (E_{NL}) as independent variables. The descriptors used to build the QSAR models for both gas and solvent phases are presented in Table 4.1.

$$\log(IC_{50}^{-1}) = -9.942 + 3.25\,\omega - 11.397\,\omega^+ - 3.71\,E_{NL} \quad (4.1)\ \text{Gas phase}$$

$$n = 16,\ r^2 = 0.706\ r_{CV}^2 = 0.430,\ SD = 1.147,\ F = 9.63,\ p < 0.05$$

$$\log(IC_{50}^{-1}) = -22.437 + 2.341\,\omega + 10.562\,\omega^+ - 5.019\,E_{NL} \quad (4.2)\ \text{Solvent phase}$$

$$n = 16,\ r^2 = 0.710\ r_{CV}^2 = 0.637,\ SD = 1.141,\ F = 9.78,\ p < 0.05$$

Here, r^2 is the square of correlation coefficient, r_{CV}^2 is the leave-one-out (LOO) cross validated squared correlation coefficient, F is the overall F-statistics for the addition of each successive term, p is the p-value using the F statistics, and SD is the standard deviation of regression. We found that the gas phase r^2 value (0.706) increases slightly (0.710) with the inclusion of solvent. However, for this case r_{CV}^2 value (0.430) increases to an acceptable value (0.637) with the change of gas phase to solvent phase indicating the importance of the solvent model.

Table 4.1: Parameters used to build the QSAR models.

Complex	$\log(IC_{50}^{-1})$		Gas phase			Solvent phase		
	A2780	A2780Cp8	ω	ω^+	E_{NL}	ω	ω^+	E_{NL}
1	-0.315	-3.807	4.134	0.868	-1.264	3.884	1.056	-0.223
2	-1.714	-4.770	3.468	0.697	-1.212	3.873	1.026	0.032
3	-3.305	-4.843	3.353	0.624	-0.616	4.107	0.920	-0.197
4	-3.660	-4.607	3.503	0.648	-0.565	4.188	0.955	-0.097
5	-1.991	-3.330	3.585	0.649	-1.523	2.956	0.834	-1.004
6	-2.179	-4.172	3.131	0.517	-1.312	2.978	0.822	-1.008
7	-5.171	-6.859	2.819	0.555	-0.501	3.078	0.605	-1.008
8	-5.120	-5.956	3.001	0.582	-0.847	3.230	0.707	-0.704
9	-2.036	-4.814	3.368	0.546	-0.907	2.820	0.663	-1.240
10	-2.220	-4.190	2.964	0.430	-0.959	2.666	0.645	-1.196
11	-4.912	-6.467	2.673	0.513	-0.469	2.907	0.610	-0.956
12	-5.121	-6.422	2.879	0.538	-0.825	3.118	0.692	-0.715
13	0.734	-1.890	2.753	0.347	-1.222	4.357	0.388	-1.607
14	-0.030	-2.111	3.072	0.602	-1.49	3.815	0.992	-0.139
15	-0.798	-2.271	2.749	0.412	-1.133	2.966	0.813	-1.209
16	-0.419	-1.319	2.608	0.344	-1.111	3.748	0.885	-0.987

In general, a regression model is significant at p-value < 0.05 using the F statistics and so these models are statistically significant.[16] However, according to the general statistical standards, a model with $r^2 > 0.80$[17] and $r_{CV}^2 > 0.60$[18] is acceptable. Therefore these QSAR equations should be further improved to become a statistically significant model. To improve r^2, one scheme was suggested by

Dietrich *et al.*[19] and Cornish-Bowden and Wong[20] in which a compound is considered as outlier if its corresponding r^2, called jackknife r^2 (r_J^2) value obtained from the regression analysis after deleting the compound, is comparatively higher than the other r_J^2 values. We applied this method to increase overall quality of the regression models. The r_J^2 values calculated in gas and solvent phases for the cell lines are presented in Table 4.2. Since the independent variables are different in both gas and solvent phases, a particular complex has quite different values of r_J^2 in gas phase from that calculated in solvent phase. Thus the outliers are different for both the phases.

Table 4.2: Jackknife results for gas and solvent phases against two cancer cell lines.

	r_J^2			
Complex	A2780		A2780Cp8	
	Gas phase	Solvent phase	Gas phase	Solvent phase
1	**0.750**	0.698	0.723	0.826
2	0.707	0.728	0.705	0.812
3	0.709	0.735	0.719	0.825
4	0.698	0.747	**0.745**	0.816
5	**0.794**	0.711	0.713	0.809
6	**0.744**	0.713	**0.748**	0.824
7	0.658	0.670	0.647	0.784
8	0.708	0.699	0.691	0.801
9	0.706	0.715	0.726	0.818
10	0.713	0.738	0.720	**0.832**
11	0.673	0.678	0.661	0.793
12	0.709	0.682	0.709	0.795
13	0.681	**0.787**	0.658	0.798
14	0.724	**0.798**	0.732	**0.936**
15	0.702	0.695	0.693	0.798
16	0.701	0.716	0.704	0.764

We observed that the complexes 1, 5, and 6 exhibited unduly high r_J^2 values (0.75, 0.794, and 0.744, respectively) in the gas phase; whereas, in the solvent phase the complexes 13, and 14 possessed higher r_J^2 values (0.787 and 0.798, respectively) and thus these complexes may be considered as outliers. However, it is seen that when complexes 5 and 6 were deleted from the data set, a significant improvement of the statistical parameters were observed compared to that obtained by deleting complexes 1 and 5.

The QSAR equations after deleting these complexes (5 and 6, in gas phase) and (13 and 14, in solvent phase) with significant statistical quality are presented in Table 4.3. We observed that r^2 values increased from 0.706 to 0.859 and r_{CV}^2 values from 0.430 to 0.748 in the gas phase. The solvent model did not show any influence for this cell line. However the solvent phase predicted r^2 and r_{CV}^2 values after applying jackknife test increased from 0.710 to 0.844 and 0.637 to 0.695, respectively.

Table 4.3: QSAR models with the statistical parameters for two cancer cell lines in gas and solvent media.

	QSAR equations	r^2	r_{CV}^2	*SD*	*F*
A2780					
Gas phase	$\log(IC_{50}^{-1}) = -12.063 + 3.852\,\omega - 12.442\,\omega^{+} - 4.855E_{NL}$	0.859	0.748	0.863	20.30
Solvent phas	$\log(IC_{50}^{-1}) = -16.087 - 0.333\omega + 14.742\,\omega^{+} - 3.566E_{NL}$	0.844	0.695	0.760	18.05
A2780Cp8					
Gas phase	$\log(IC_{50}^{-1}) = -7.593 + 1.493\,\omega - 9.743\,\omega^{+} - 4.083\,E_{NL}$	0.794	0.568	0.905	12.91
Solvent phase	$\log(IC_{50}^{-1}) = -23.184 + 2.129\,\omega + 9.717\,\omega^{+} - 5.11\,E_{NL}$	0.954	0.908	0.403	69.66

4.4.2 QSAR Analysis on A2780Cp8 Cell Line

Multi-linear regression analysis between $\log(IC_{50}^{-1})$ of platinum complexes (1–16, Figure 4.1) against A2780Cp8 cell line and the combination of three DFT derived descriptors yielded the following QSAR equations:

$$\log(IC_{50}^{-1}) = -8.510 + 2.313\,\omega - 10.984\,\omega^{+} - 3.13\,E_{NL} \quad (4.3) \text{ Gas phase}$$

$$n = 16,\ r^2 = 0.702,\ r_{CV}^2 = 0.450,\ SD = 1.00,\ F = 9.43,\ p < 0.05$$

$$\log(IC_{50}^{-1}) = -22.766 + 2.06\,\omega + 10.016\,\omega^{+} - 4.846\,E_{NL} \quad (4.4) \text{ Solvent phase}$$

$$n = 16,\ r^2 = 0.813,\ r_{CV}^2 = 0.595,\ SD = 0.796,\ F = 17.39,\ p < 0.05$$

The influence of solvent effect was very much prominent for this cell line. The r^2 and r_{CV}^2 values (0.702 and 0.45, respectively) obtained in the gas phase increased to 0.813 and 0.595, respectively, with the inclusion of the solvent. Although the model in the solvent medium displayed acceptable statistical quality revealing the importance of the descriptors in the determination of biological activity of platinum complexes, the jackknife test may provide more insight in building more significant models for the cell line in both the media.

The complexes 4 and 6 with higher r_J^2 values (0.745 and 0.748, respectively) in the gas phase and complexes 10 and 14 indicating higher r_J^2 values (0.832 and 0.936, respectively) in the solvent phase could be considered as outliers (Table 4.2). The QSAR equations obtained after deleting these complexes are given in Table 4.3 along with statistically significant parameters. Importantly, for this cell line, the r^2 value 0.794 obtained after applying the jackknife test increased to 0.954 and r_{CV}^2 value increased from 0.568 to a very acceptable value of 0.908 in solvent medium demonstrating the importance of the selected descriptors in the determination of $\log(IC_{50}^{-1})$ values of platinum complexes. Autocorrelation coefficients among the descriptors of QSAR models are reasonable. We found that in gas phase E_{NL} had very low correlations with ω and ω^{+} (< 0.3) for both the cell lines. Similarly, in solvent phase, ω had low correlations with E_{NL} and ω^{+} (< 0.5). We found slightly higher autocorrelation between ω and ω^{+} in gas phase and E_{NL} and ω^{+} in solvent phase. However, models having descriptors with autocorrelation of about 0.8 have been reported for QSAR analyses.[17]

In platinum drug-DNA binding, the DNA molecule acts as an electron donor whereas the complex is an electron acceptor and the mechanism involves the nucleophilic attack at Pt atom. In this type of interaction E_{LUMO} and E_{NL} play an important role. The lower values of these parameters increase the capability of the molecules to accept electrons from DNA making the system stable. We found that

the coefficients of E_{NL} in all the QSAR equations (Table 4.3) are negative suggesting highly favorable intermolecular interactions between DNA molecule and the complex and an enhanced cytotoxic activity for the complex. The coefficients of other two independent factors (ω and ω^+), however were not consistent in all equations. Importantly, the most significant model (r^2 =0.954 and r^2_{CV} =0.908) had positive coefficients for ω and ω^+. Thus, increasing their values can improve the anticancer activity. Although, all QSAR models are statistically significant, we found solvent phase derived model with r^2 =0.954 and r^2_{CV} =0.908, and gas phase derived model with r^2 =0.859 and r^2_{CV} =0.748 as the best models. The standard errors of regression coefficients (S_β) for two cancer cell lines in gas and solvent phases are calculated and provided in Table 4.4. We found that the best two models have lower values of S_β than that of other two models.

Table 4.4: Standard errors of regression coefficients (S_β) for 14 platinum complexes in QSAR models.

Cell line	S_β	
	Gas phase	Solvent phase
A2780	2.321	3.473
	1.357	1.098
	4.006	2.531
	0.841	0.992
A2780Cp8	2.432	1.359
	1.531	0.234
	4.347	0.953
	0.939	0.382

The correlation plots between experimental and calculated $\log(IC_{50}^{-1})$ values of the platinum complexes derived from these two more significant QSAR models are

shown in Figure 4.2 which indicates that these descriptors can be effectively used in the prediction of cytotoxicity of platinum complexes.

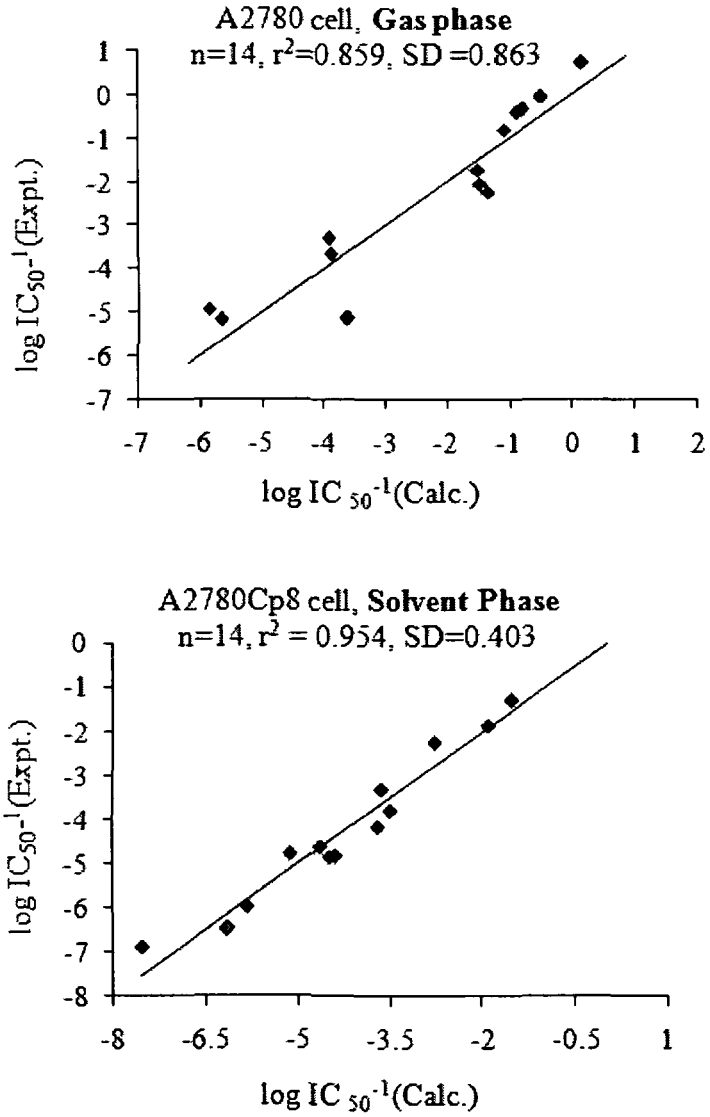


Figure 4.2: Plots of experimental versus calculated values of cytotoxicity $\log(IC_{50}^{-1})$ for two best models.

4.4.3 QSPR Analysis

We carried out QSPR studies for analyzing the $\log P_{o/w}$ values of platinum complexes (1–24, Figure 4.1 and Figure 4.3) for 0% (extrapolated), 20%, 30%, 40%, and 50% MeOH. These $\log P_{o/w}$ values were estimated by reversed-phase high performance liquid chromatography (RP-HPLC) technique.[6]

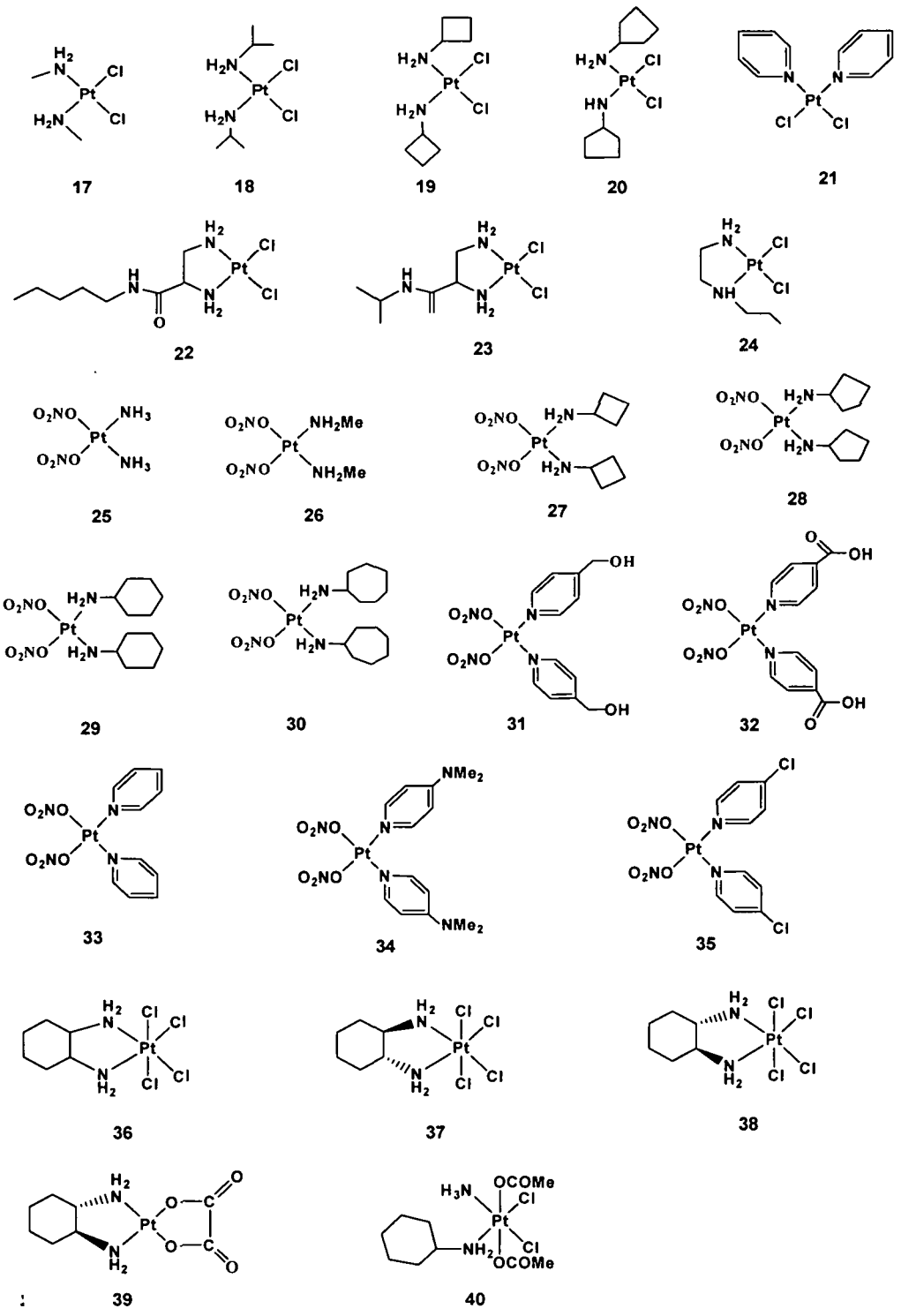


Figure 4.3: Sketch of the platinum complexes used to build QSPR models.

Multi-linear regression analyses were performed using the experimental $\log P_{o/w}$ values for 0-50% MeOH as a dependent variable and combination of four

descriptors, namely ω, ω^+, MR_{CL}, and SA of the complexes as independent variables in gas and solvent models (Table 4.5).

Table 4.5: Parameters used to build the QSPR models for 24 platinum complexes in both gas and solvent phases.

	$\log P_{o/w}$ (% MeOH)					Gas phase				Solvent phase			
	0	20	30	40	50	ω	ω^+	MR	SA	ω	ω^+	MR	SA
1	-2.27	-2.28	-2.08	-2.15	-2.09	4.134	0.868	4.8	142.2	3.884	1.056	4.8	144.64
2	-2.16	-2.20	-2.04	-2.08	-2.19	3.468	0.697	12.62	167.56	3.873	1.056	12.62	168.86
3	-0.85	-0.87	-0.90	-0.98	-0.97	3.353	0.624	32.21	242.64	4.107	0.920	32.21	243.6
4	-1.23	-1.24	-1.22	-1.32	-1.27	3.503	0.648	30.24	228.14	4.188	0.955	30.24	228.66
5	-2.32	-2.35	-2.13	-2.19	-2.20	3.585	0.649	4.8	172.88	2.956	0.834	4.8	176.75
6	-2.19	-2.25	-2.04	-2.15	-2.23	3.131	0.517	12.62	197.29	2.978	0.822	12.62	197.9
7	-1.17	-1.16	-1.26	-1.31	-1.27	2.819	0.555	32.21	275.92	3.078	0.605	32.21	277.82
8	-1.47	-1.44	-1.47	-1.46	-1.37	3.001	0.582	30.24	260.17	3.230	0.707	30.24	259.95
9	-1.63	-1.69	-1.60	-1.72	-1.80	3.368	0.546	4.8	217.47	2.820	0.663	4.8	223.54
10	-1.70	-1.70	-1.66	-1.67	-1.64	2.964	0.430	12.62	241.82	2.666	0.645	12.62	243.55
11	-0.47	-0.49	-0.63	-0.78	-0.75	2.673	0.513	32.21	320.62	2.907	0.610	32.21	323.19
12	-0.79	-0.81	-0.95	-1.05	-1.06	2.879	0.538	30.24	304.23	3.118	0.692	30.24	304.97
13	-1.39	-1.41	-1.37	-1.46	-1.43	2.753	0.347	28.71	237.77	4.357	0.388	28.71	237.18
14	-1.40	-1.41	-1.34	-1.41	-1.34	3.072	0.602	28.71	226.24	3.815	0.992	28.71	228.19
15	-1.37	-1.40	-1.35	-1.44	-1.43	2.749	0.412	28.71	253.08	2.966	0.813	28.71	257.17
16	-0.85	-0.88	-0.92	-1.02	-1.04	2.608	0.344	28.71	301.85	3.748	0.885	28.71	301.94
17	-1.94	-1.90	-1.93	-1.86	-1.79	3.654	0.735	14.59	187.72	3.289	0.783	14.59	190.17
18	-0.61	-0.54	-0.91	-0.65	-0.72	3.382	0.643	32.13	253.83	4.911	0.783	32.13	271.44
19	0.09	0.06	0.06	0.00	-0.05	3.500	0.668	38.22	283.27	5.621	1.360	38.22	283.88
20	1.06	1.07	1.14	1.16	1.19	3.411	0.645	47.42	315.4	5.431	1.350	47.42	315.18
21	-0.04	-0.18	-0.41	-0.78	-1.06	5.425	0.705	13.69	263.8	4.876	0.880	13.69	260.75
22	-0.04	-0.06	0.12	0.12	0.13	3.778	0.362	40.73	310.73	4.227	0.980	40.73	312.54
23	-0.46	-0.55	-0.64	-0.86	-1.01	3.934	0.377	31.42	267.82	4.137	1.170	31.42	268.01
24	-2.12	-2.25	-2.22	-2.38	-2.55	3.265	0.580	23.81	214.73	3.169	0.815	23.81	209.75

The QSPR equations obtained in both gas and solvent phases with r^2, r^2_{CV}, *SD*, and *F*-values are listed in Table 4.6.

Table 4.6: QSPR models for 24 platinum complexes with the statistical parameters.

MeOH %	QSPR equations	r^2	r^2_{CV}	*SD*	*F*
Gas phase					
0	$\log P_{o/w} = -7.253 + 0.559\,\omega + 0.812\,\omega^+ + 0.025\,MR_{CL} + 0.012\,SA$	0.914	0.798	0.274	50.63
20	$\log P_{o/w} = -7.234 + 0.512\,\omega + 0.986\,\omega^+ + 0.026\,MR_{CL} + 0.012\,SA$	0.908	0.845	0.284	47.10
30	$\log P_{o/w} = -6.473 + 0.502\,\omega + 0.648\,\omega^+ + 0.029\,MR_{CL} + 0.010\,SA$	0.86	0.734	0.329	29.17
40	$\log P_{o/w} = -6.304 + 0.414\,\omega + 0.925\,\omega^+ + 0.033\,MR_{CL} + 0.009\,SA$	0.833	0.656	0.362	23.71
50	$\log P_{o/w} = -6.071 + 0.312\,\omega + 1.140\,\omega^+ + 0.035\,MR_{CL} + 0.009\,SA$	0.812	0.602	0.388	20.53
Solvent phase					
0	$\log P_{o/w} = -6.859 + 0.449\,\omega + 0.651\,\omega^+ - 0.013\,MR_{CL} + 0.015\,SA$	0.952	0.914	0.203	95.33
20	$\log P_{o/w} = -6.802 + 0.449\,\omega + 0.591\,\omega^+ - 0.010\,MR_{CL} + 0.015\,SA$	0.951	0.917	0.206	93.63
30	$\log P_{o/w} = -6.375 + 0.363\,\omega + 0.894\,\omega^+ - 0.008\,MR_{CL} + 0.013\,SA$	0.927	0.867	0.236	60.77
40	$\log P_{o/w} = -6.232 + 0.369\,\omega + 0.790\,\omega^+ - 0.001\,MR_{CL} + 0.012\,SA$	0.916	0.855	0.257	51.82
50	$\log P_{o/w} = -5.947 + 0.321\,\omega + 0.743\,\omega^+ + 0.005\,MR_{CL} + 0.011\,SA$	0.882	0.795	0.307	35.68

As expected the solvent phase models displayed higher predictive power with r^2_{CV} values ranging from 0.914 for 0% MeOH to 0.795 for 50% MeOH. The statistical significance of the models of $\log P_{o/w}$ in the solvent phase was similar for 0% and 20% MeOH but different in the gas phase. Although predictability of the models decreases for higher concentration, the solvent phase derived QSPR equations for 30%–50% MeOH could predict the partition coefficient with the r^2_{CV} values in the range 0.867–0.796, respectively. The standard errors of regression

coefficients for QSPR models in all concentrations are lower in solvent phase than that in gas phase (Table 4.7). The sign of all descriptors in gas phase are in consistent with that obtained in solvent phase, except for MR_{CL}. But we found positive correlation of $\log P_{o/w}$ values of the complexes at all concentrations of MeOH with MR_{CL} while taking it as a single descriptor. The steric factor of the amine carrier ligands can be expressed by MR_{CL}. Thus, greater the steric effect of carrier amine ligands, greater might be the hydrophobicity of the complexes, in agreement with the observations by Platts *et al.*[6] Importantly, the complexes (19, 20, and 22) with bulkiest amine carrier ligands and higher values of molar refractivity (38.22, 47.42, and 40.73, respectively) exhibited higher values of $\log P_{o/w}$.

Table 4.7: Standard errors of regression coefficients (S_β) for 24 platinum complexes in QSPR models for five concentrations of MeOH.

	S_β				
MeOH %	0	20	30	40	50
Gas phase	0.628	0.651	0.755	0.831	0.890
	0.117	0.121	0.141	0.155	0.166
	0.589	0.611	0.708	0.779	0.835
	0.008	0.008	0.010	0.010	0.010
	0.002	0.002	0.003	0.003	0.003
Solvent phase	0.374	0.379	0.435	0.472	0.564
	0.074	0.075	0.087	0.094	0.112
	0.255	0.258	0.296	0.322	0.384
	0.007	0.007	0.008	0.009	0.010
	0.001	0.001	0.002	0.002	0.002

The plots between experimental and calculated values of $\log P_{o/w}$ for 0% MeOH predicted by gas and solvent phases presented in Figure 4.4 suggest that the selected descriptors can be effectively used in the determination of $\log P_{o/w}$ of the complexes.

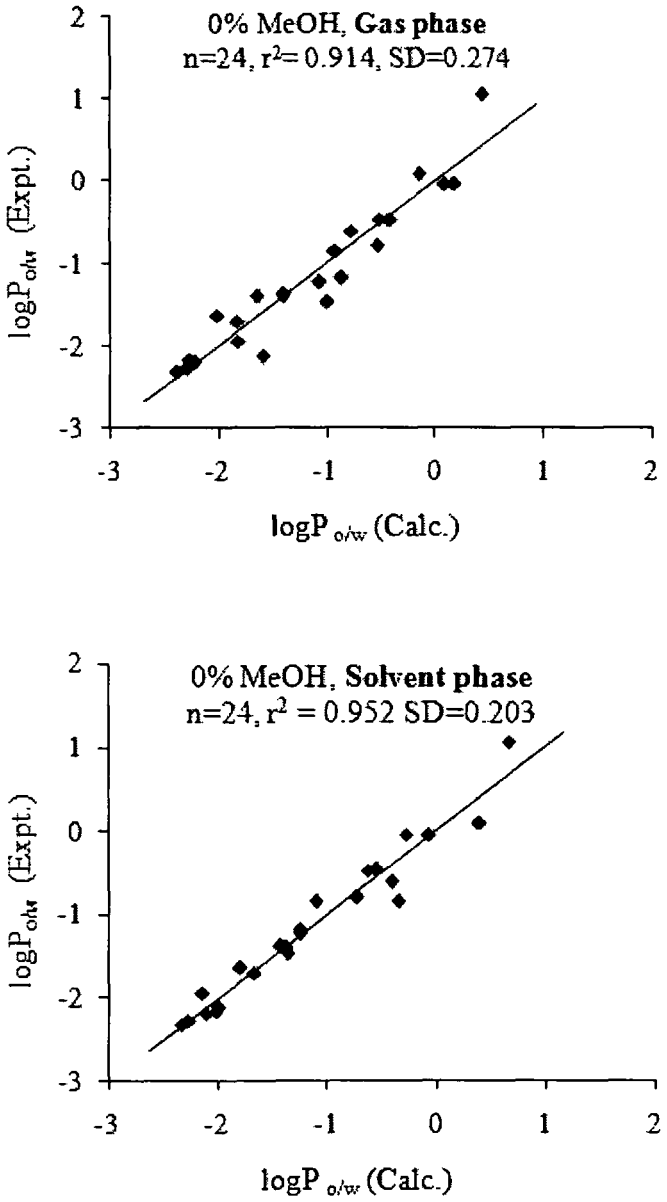


Figure 4.4: Experimental versus calculated log$P_{o/w}$ values of 24 platinum complexes at 0% MeOH in gas and solvent media.

We have also carried out QSPR analysis for an additional set of 20 complexes, log$P_{o/w}$ values of which were calculated by standard shake-flask method. The log$P_{o/w}$ values of complexes 1, 9, 13, 17, 19, 20, 21 and 25 to 40 (Figure 4.1 and Figure 4.3) were taken from results reported by Screnci *et al.*[5] and Souchard *et al.*[21] (Both these papers reported values for 1). The complexes 1 (from reference 5), 9, 13, 20, and 25-40 (a set of 20 compounds) were considered as a training set and the other four compounds (1, 17, 19, and 21) were treated as a test set. Table 4.8 lists

the values of $\log P_{o/w}$ and other descriptors derived from gas and solvent phases for the training set.

Table 4.8: Parameters used to build the QSPR models for 20 platinum complexes.

	$\log P_{o/w}$	Gas phase				Solvent phase			
Complex		ω	ω^+	MR_{CL}	SA	ω	ω^+	MR_{CL}	SA
1	-2.53	4.134	0.868	4.8	142.2	3.884	1.056	4.8	144.64
9	-2.30	3.368	0.546	4.8	217.47	2.820	0.663	4.8	223.54
13	-1.65	2.753	0.347	28.71	237.77	4.357	0.388	28.71	237.18
20	0.81	3.411	0.645	47.42	315.4	5.431	1.350	47.42	315.18
25	-3.36	5.283	0.718	4.8	183.13	7.948	0.540	4.8	224.13
26	-3.28	5.263	0.511	14.59	224.72	6.946	0.392	14.59	225.07
27	-1.71	4.727	0.397	38.22	323.12	6.967	0.385	38.22	318.91
28	-1.14	6.967	0.507	47.42	351.86	6.934	0.362	47.42	353.65
29	-0.91	6.649	0.390	56.62	386.53	8.755	0.359	56.62	397.07
30	-0.35	7.759	0.698	62.62	382.13	7.957	0.607	62.62	378.95
31	-2.13	8.821	0.547	26.46	372.26	8.083	0.225	26.46	375.7
32	-1.41	8.496	0.509	30.43	418.04	9.617	0.326	30.43	419.56
33	-1.59	6.447	0.621	13.69	308.75	7.304	0.727	13.69	315.54
34	-0.83	6.596	0.387	40.49	417.54	7.547	0.237	40.49	425.23
35	-1.06	4.818	0.384	22.68	346.64	6.933	0.643	22.68	349.61
36	-1.17	7.430	0.952	28.71	276.97	9.835	1.321	28.71	278.99
37	-1.18	7.430	0.952	28.71	276.97	9.835	1.321	28.71	278.99
38	-1.03	7.430	0.952	28.71	276.97	9.835	1.321	28.71	278.99
39	-1.59	2.753	0.347	28.71	237.77	4.357	0.388	28.71	237.18
40	-0.16	7.191	0.683	30.71	371.37	8.942	1.096	30.71	374.59

The multi-linear regression analysis between $\log P_{o/w}$ values of these 20 complexes and four descriptors yielded the QSPR equations as shown below.

$\log P_{o/w} = -6.849 - 0.623\,\omega + 4.864\,\omega^{+} + 0.014\,MR_{CL} + 0.018\,SA$ (4.5) Gas phase

$n = 20,\ r^2 = 0.946,\ r^2_{CV} = 0.913\ \ SD = 0.251\ \ F = 65.79$

$\log P_{o/w} = -4.947 - 0.230\,\omega + 1.884\,\omega^{+} + 0.027\,MR_{CL} + 0.010\,SA$ (4.6) Solvent Phase

$n = 20,\ r^2 = 0.955,\ r^2_{CV} = 0.920\ \ SD = 0.228\ \ F = 80.67$

As expected, the solvent phase played an important role in improving the statistical quality of the model. The predicted $\log P_{o/w}$ values of the compounds in the test set are presented in Table 4.9.

Table 4.9: Experimental and predicted $\log P_{o/w}$ values of 4 complexes in the test set.

	$\log P_{o/w}$ [a]	Gas phase		Solvent phase	
Complex		$\log P_{o/w}$ [b]	Residual [c]	$\log P_{o/w}$ [d]	Residual [c]
1	-2.19	-2.443	0.253	-2.264	0.0742
17	-1.68	-1.785	0.105	-1.917	0.237
19	0.36	0.142	0.217	0.216	0.143
21	-0.32	-1.610	1.290	-1.412	1.092

a) Experimental $\log P_{o/w}$ values obtained from Ref. 5 and Ref. 21
b) Predicted $\log P_{o/w}$ values calculated using Eq. (4.5)
c) Difference between the experimental and calculated values of $\log P_{o/w}$
d) Predicted $\log P_{o/w}$ values calculated using Eq. (4.6)

The training set and test set were then combined and multi-linear regression analysis was performed on this data set of 24 complexes. The QSPR equations obtained in gas and solvent phases are:

$\log P_{o/w}$ = -7.057 - 0.677 ω + 5.320 ω^+ + 0.008 MR_{CL} + 0.020 SA (4.7) Gas phase

n =24, r^2 = 0.895, r^2_{CV} = 0.842 SD = 0.393 F = 32.05

$\log P_{o/w}$ = - 4.917 - 0.275 ω + 2.031 ω^+ + 0.022 MR_{CL} + 0.011 SA (4.8) Solvent Phase

n =24, r^2 = 0.922, r^2_{CV} = 0.882 SD = 0.339 F = 44.32

The sign of coefficients of all descriptors are same for both training set and data set in gas phase as well as in solvent phase. However, their signs for ω and MR_{CL} are different from those obtained in QSPR models for five different concentrations of MeOH. This inconsistency may be due to their values of $\log P_{o/w}$ calculated from different experimental techniques. The standard errors of regression coefficients for QSPR models of both training and data set are lower in solvent phase than that in gas phase (Table 4.10). Together, these results demonstrate that the four descriptors (ω, ω^+, MR_{CL}, and SA) can be satisfactorily used in the prediction of hydrophobicity of platinum complexes and the solvent model derived descriptors provide a better correlation.

Table 4.10: Standard errors of regression coefficients (S_β) for a training set of 20 complexes and data set of 24 platinum complexes in QSPR models

	S_β	
	Gas phase	Solvent phase
Training set	0.400	0.260
	0.058	0.031
	0.433	0.144
	0.006	0.004
	0.002	0.001
Data set	0.603	0.354
	0.081	0.042
	0.620	0.192
	0.007	0.006
	0.003	0.002

The correlation plots between experimental and calculated log$P_{o/w}$ values of the data set in gas and solvent media presented in Figure 4.5 displayed a good correlation among them. The correlation in the solvent phase was better than that of gas phase as expected. This indicates that the selected parameters can predict the hydrophobicity of platinum complexes with greater predictability in the solvent phase.

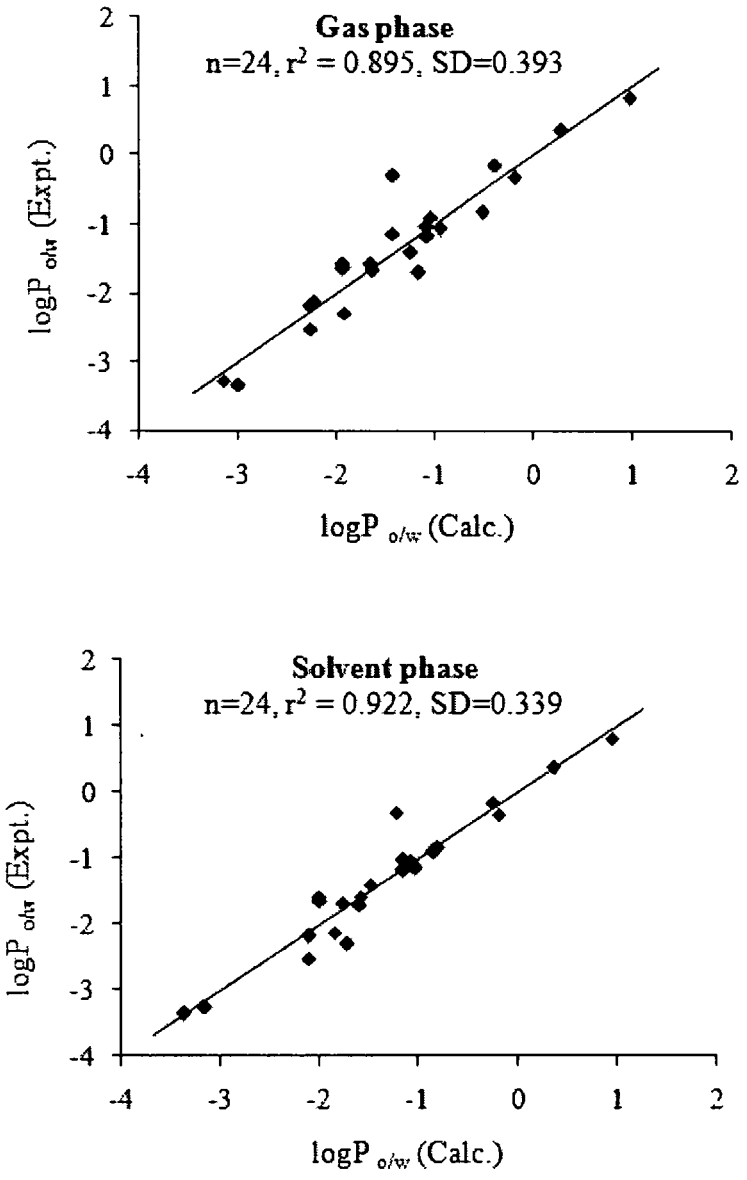


Figure 4.5: Correlation plots between experimental and calculated log$P_{o/w}$ values of the data set obtained in gas and solvent phases.

4.5 Conclusions

The results presented in this work clearly show that the QSAR approach with three parameters i.e., ω, ω^+, and E_{NL} is capable of predicting $\log(IC_{50}^{-1})$ of *cis*-platinum complexes. The jackknife test applied on the QSAR studies improved the statistical quality of the models in both gas and solvent phases. The inclusion of solvent medium increases the correlation coefficient ($r^2 = 0.954$) and cross-validated squared correlation ($r^2_{CV} = 0.908$) for the A2780Cp8 cancer cell line suggesting the importance of solvent effect and significance of the selected descriptors. The QSPR equations modeled by ω, ω^+, MR_{CL}, and SA parameters against five different concentrations of MeOH (0%–50%) are capable of predicting $\log P_{o/w}$ values of 24 platinum complexes with r^2 values in the range of 0.914–0.812 (gas phase) and 0.952–0.882 (solvent phase) and r^2_{CV} values in the range of 0.798–0.602 (gas phase) and 0.914–0.795 (solvent phase), respectively. The solvent effect influences the QSPR model developed for 20 platinum complexes (training set) where calculated $\log P_{o/w}$ values are in close proximity to their experimental values with $r^2 = 0.946$ (0.955) and $r^2_{CV} = 0.913$ (0.920) for gas (solvent) media. The predicted $\log P_{o/w}$ values of 4 complexes in test set derived from the models are near to their corresponding experimental values, indicating significance of the selected descriptors in determination of hydrophobicity of platinum complexes. QSPR models for the data set also show good statistic qualities in both gas and solvent phases. Thus the preset study indicates the importance of the selected parameters as well as solvent effect in the QSAR and QSPR analyses of *cis*-platinum complexes.

References

(1) Jakupec, M.A.; Galanski, M.; Keppler, B.K. Tumour inhibiting platinum complexes state of the art and future perspectives. *Rev. Physiol. Biochem. Pharmacol.* **146,** 1–53 (2003)

(2) Giaccone, G. Clinical perspectives on platinum resistance. *Drugs* **59**, 9–17 (2000)

(3) Weiss, R.B.; Christian, M.C. New cisplatin analogues in development. A review. *Drugs* **46**, 360–377 (1993)

(4) Monti, E. *et al* Cytotoxicity of *cis*-platinum(II) conjugate models. The effect of chelating arms and leaving groups on cytotoxicity: A quantitative structure-activity relationship approach. *J. Med. Chem.* **48**, 857–866 (2005)

(5) Screnci, D. *et al* Relationships between hydrophobicity, reactivity, accumulation and peripheral nerve toxicity of a series of platinum drugs. *British. J. Cancer* **82**, 966–972 (2000)

(6) Platts, J.A. *et al* The RP-HPLC measurement and QSPR analysis of $\log P_{o/w}$ values of several Pt(II) complexes. *J. Inorg. Biochem.* **100**, 1199–1207 (2006)

(7) Platts, J.A.; Hibbs, D.E.; Hambley, T.W.; Hall, M.D. Calculation of the hydrophobicity of platinum drugs. *J. Med. Chem.* **44**, 472–474 (2001)

(8) Costa, L.A.S.; Rocha, W.R.; De Almeida, W.B.; Santos, H.F.D. Linear free energy relationship for 4-substituted (o-phenylenediamine)platinum (II) dichloride derivatives using quantum mechanical descriptors. *J. Inorg. Biochem.* **99**, 575–583 (2005)

(9) Hirshfeld, F.L. The theoretical tool used is the minimum entropy deficiency principle minimum missing information principle. *Theor. Chim. Acta.* **44**, 129–138 (1977)

(10) HyperChem Release 7 (2002), Hypercube; http://www.hyper.com

(11) MATLAB; The Math Works, Inc.: Natick, USA, 1999.

(12) Penrose, R. A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.* **51**, 406–413 (1955)

(13) Milburn, G.H.W.; Truter, M.R. The crystal structures of cis- and trans-dichlorodiammineplatinum (II). *J. Chem. Soc. A* **11**, 1609–1616 (1966)

(14) Beagley, B. *et al* The crystal and molecular structure of *cis*-diammine-1,1-cyclobutanedicarboxoplatinum(II) [*cis*-$Pt(NH_3)_2CBDCA$]. Dynamic puckering of the cyclobutane ring. *J. Mol. Struct.* **130**, 97–102 (1985)

(15) Bruck, M.A.; Bau, R.; Noji, M.; Inagaki, K.; Kidani, Y. The Crystal structures and absolute configurations of the anti-tumor complexes Pt(oxalato)(1R,2R-cyclohexanediamine) and Pt(malonato)(1R,2R-cyclohexanediamine). *Inorg. Chim. Acta.* **92**, 279–284 (1984)

(16) Cho, D.H.; Lee, S.K.; Kim, B.T.; No, K.T. Quantitative structure-activity relationship (QSAR) study of new fluorovinyloxyacetamides. *Bull. Korean. Chem. Soc.* **22**, 388–394 (2001)

(17) Yao, S.W. *et al* Synthesis and QSAR study of the anticancer activity of some novel indane carbocyclic nucleosides. *Bioorg. Med. Chem.* **11**, 4999–5006 (2003)

(18) Wold, S. Validation of QSARs. *Quant. Struct-Act. Relat.* **10**, 191–193 (1991)

(19) Dietrich, S.W.; Dreyer, N.D.; Hansch, C.; Bentley, D.L. Confidence interval estimators for parameters associated with quantitative structure-activity relationships. *J. Med. Chem.* **23**, 1201–1205 (1980)

(20) Cornish-Bowden, A.; Wong, J.T. Evaluation of rate constants for enzyme-catalysed reactions by the jackknife technique. Application to liver alcohol dehydrogenase. *Biochem. J.* **175**, 969–976 (1978)

(21) Souchard, J.P.; Ha, T.T.B.; Cros, S.; Johnson, N.P. Hydrophobicity parameters for platinum complexes. *J. Med. Chem.* **34**, 863–864 (1991)

Chapter 5

QSAR Study and Molecular Design of Nucleoside Analogues as Potent Anticancer Agents

We have performed multiple linear regression analyses to build QSAR models for nucleoside analogues using DFT and MM+ based descriptors in both gas and solvent phases. The QSAR models for carbocyclic analogues of nucleosides against murine leukemia cell line (L1210/0) and human T-lymphocyte cell lines (Molt4/C8 and CEM/0) explain more than 90% of the variances in the activity data along with higher values of r^2_{CV} (>0.86). Inclusion of solvent medium increases the correlation of each descriptor with activity. Based on the key features responsible for anticancer activity, 10 new compounds with rather high anticancer activity have been theoretically designed. Anticancer activities of an additional set of 20 nucleoside analogues have also been modeled by the same descriptors and found their predicted values to be in good agreement with the experimental values. [Sarmah and Deka, *J. Mol. Model.* **16**, 411–418 (2010)]

5.1 Introduction

The nucleoside analogues were among the first anticancer agents to be introduced for the medical treatment of cancer.[1,2] Certain nucleosides as anticancer drugs have been used clinically for many years and recently a large number of nucleoside analogues with anticancer activity have been designed and synthesized.[3-5] Of particular interest are carbocyclic nucleoside analogues in which the core of the pseudo-sugar moiety is a cyclopentane ring. Carbocyclic nucleosides act as chain terminators when incorporate into the growing DNA chain, that is, they do not allow further chain elongation and thought to inhibit the DNA synthesis. As a result, they can possess antiviral, antineoplastic activity or other potentially biological properties.[6,7] Over the years research has been engaged in the synthesis, evaluation of activity and design of carbocyclic nucleoside analogues in an effort to develop potential anticancer or antiviral agents.[8-14] Several derivatives of carbocyclic nucleosides, Carbovir and Abacavir[15,16], potent anticancer agents, have been recently synthesized in which the cyclopentene ring is replaced by an indane system and assayed on different cancer cell lines.[12] Santana *et al.*[17] synthesized a series of 1′,2′-*cis*-disubstituted 8-azapurine-based carbocyclic analogues of nucleosides and evaluated their antitumoral activities against L1210, Molt4/C8, and CEM/0 cell lines.

The anticancer mechanism of nucleoside analogues has not been properly clarified yet. The steric, lipophilic, and electronic properties of molecules play important roles in the drug–DNA interaction and have been used successfully in many QSAR studies and drug designs.[18-20] Among different descriptors for describing the electronic properties of molecules, the quantum chemical descriptors based on DFT and semi-empirical methods have been found useful in several QSAR studies.[21,22] In particular, net atomic charges, HOMO-LUMO energies, frontier orbital electron densities, and superdelocalizabilities have shown to correlate with various biological activities.[23] Yao *et al.*[12] reported QSAR models of carbocyclic nucleosides based on steric and electronic properties of the molecules which reproduced well the experimental data against three cancer cell lines. Helguera *et al.*[5] performed QSAR modelling of a large dataset of nucleosides to probe anticancer activity against L1210/0 cancer cell line using linear discriminant analysis along with 2D-molecular descriptors. Their obtained QSAR models show very good

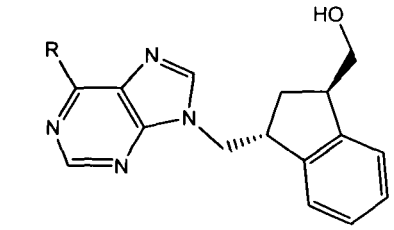
overall accuracy and predictability on external data set. DFT based reactivity descriptors have been shown to be effective in predicting anticancer activity of platinum complexes form our previous studies in Chapters 3 and 4. The goal of the present chapter is to extend this investigation and to perform QSAR analysis of a series of nucleoside analogues for designing some new compounds.

5.2 Computational Details

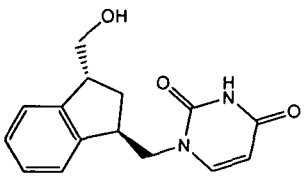
Structures of all nucleoside analogues are presented in Figure 5.1. Full geometry optimizations of these compounds without symmetry constraints along with frequency calculations were carried out using DMol^3 program at BLYP/DNP level. The reactivity descriptors electrophilicity index (ω), chemical potential (μ), and global hardness (η) were calculated for all systems. To investigate the solvent effect the Conductor-like Screening Model (COSMO) with dielectric constant of 78.4 was adopted. In addition, the molar refractivity (*MR*), van der Waals surface area (*SA*), volume (*V*), mass (*M*), and lipophilicity index (*logP*) for whole molecule were calculated from the MM+ computations with HyperChem software.

5.3 QSAR Modelling

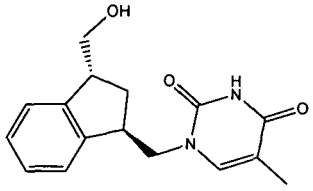
The anticancer activity data of compounds (1-14, Figure 5.1) against the murine leukemia cell line (L1210/0) and human T-lymphocyte cell lines (Molt4/C8 and CEM/0) were collected from the literature.[12] All these activities calibrated to the logarithmic log(IC_{50}^{-1}) values are listed in Table 5.1. The analyses were performed in both gas and solvent media. From the results of DFT calculations different descriptors, *viz.*, energy of HOMO (E_{HOMO}), energy of LUMO (E_{LUMO}), energy of the next lowest unoccupied molecular orbital (E_{NL}), energy difference between LUMO and HOMO ($\Delta_{L\text{-}H}$), dipole moments, electrophilicity (ω), hardness (η), and chemical potential (μ) were selected for QSAR analysis. In addition, the molecular mechanics parameters such as molar refractivity (*MR*), van der Waals surface area (*SA*), molecular volume (*V*), mass (*M*), and hydrophobicity (*logP*) of the compounds were also selected. The descriptors with greater correlation to log(IC_{50}^{-1}) with smaller autocorrelation were selected out to perform the stepwise multi-linear regression. The predictive power of the models was validated by using "leave one out" (LOO) cross-validation method.



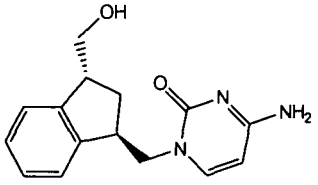
Compound	R
1	NH_2
2	$NHCH_2CH_2CH_3$
3	$NHCH(CH_3)_2$
4	$NHCH(CH_2)_2$
5	OH
6	OCH_3
7	OCH_2CH_3
8	C_6H_5
9	$C_6H_4CH_3$
10	C_6H_4Cl
11	$C_6H_4OCH_3$



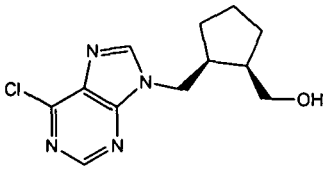
12



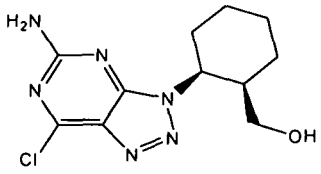
13



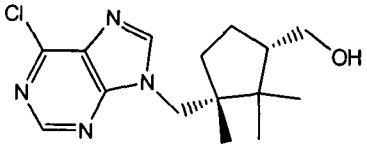
14



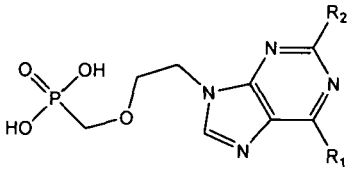
15



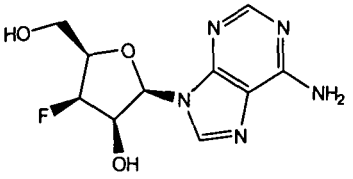
16



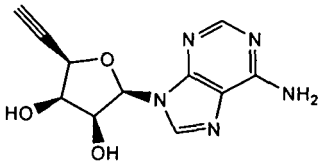
17



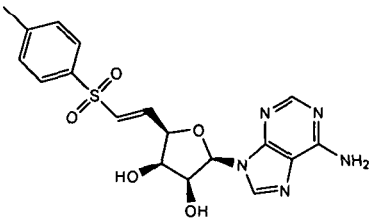
Compound	R_1	R_2
18	NH_2	NH_2
19	NH_2	C_3H_6N
20	O	NH_2
21	NH_2	H



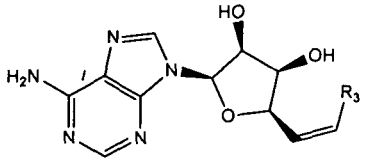
22



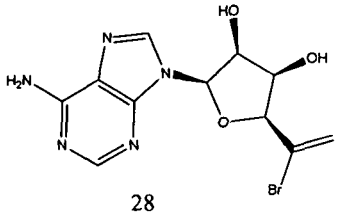
23



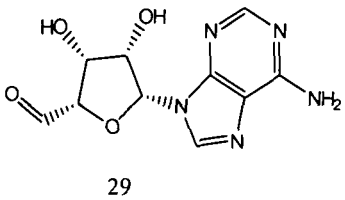
24



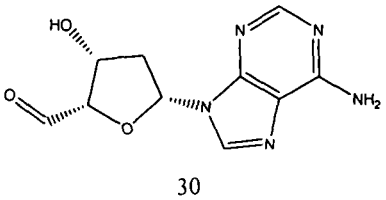
Compound	R_3
25	Br
26	Cl
27	I



28



29



30

Figure 5.1: Sketch of the nucleoside analogues used to build QSAR models.

5.4 Results and Discussion

We found electrophilicity (ω), energy of the next LUMO (E_{NL}), and van der Waals surface area (SA) as most relevant descriptors for modelling inhibitory activity of carbocyclic nucleosides (1-14, Figure 5.1), values of which in both gas and solvent phases are presented in Table 5.1.

Table 5.1: Parameters used to build the QSAR models for 14 carbocyclic nucleosides in gas and solvent phases against three cancer cell lines.

	$\log(IC_{50}^{-1})$			Gas phase			Solvent phase		
	L1210/0	Molt4/C8	CEM/0	ω	E_{NL}	SA	ω	E_{NL}	SA
1	-2.53	-2.44	-2.44	3.129	-1.146	313.98	3.681	-1.167	313.69
2	-2.32	-2.04	-2.16	2.883	-1.119	375.29	3.103	-1.155	375.64
3	-2.20	-1.85	-1.93	2.884	-1.125	373.58	3.525	-1.148	374.37
4	-2.40	-2.24	-2.31	2.973	-1.129	366.61	3.202	-1.172	366.85
5	-2.42	-2.16	-2.21	3.843	-1.345	307.53	3.947	-1.387	307.49
6	-2.54	-2.33	-2.37	3.598	-1.216	330	3.848	-1.285	330.73
7	-2.33	-1.79	-1.79	3.519	-1.204	351.67	3.802	-1.27	350.04
8	-1.45	-0.980	-0.954	5.469	-1.410	376.77	5.851	-1.50	377.63
9	-1.19	-0.636	-0.578	5.271	-1.351	397.44	5.708	-1.469	397.62
10	-0.927	-0.317	-0.208	5.967	-1.576	393.70	5.906	-1.622	392.90
11	-1.25	-0.919	-0.845	5.042	-1.311	406.46	5.601	-1.477	406.46
12	-2.87	-2.87	-2.87	4.558	-1.282	288.80	4.321	-1.143	289.38
13	-2.59	-2.45	-2.45	4.259	-1.243	309.59	4.092	-1.132	309.43
14	-2.87	-2.87	-2.87	3.784	-1.042	320.46	3.680	-1.100	295.67

The best fit QSAR equations with absolute values of statistical parameters for these molecules against L1210/0, Molt4/C8, and CEM/0 cell lines in both gas and solvent phases are represented in Table 5.2 and Table 5.3, respectively. The models were calculated by considering the inhibitory activity ($\log(IC_{50}^{-1})$) as a dependent variable and possible combination of other descriptors such as ω, S, and E_{NL} as independent variables. The quality of calculated models was measured by the square of correlation coefficient, r^2, the leave-one-out (LOO) cross-validated squared

correlation coefficient, r^2_{CV}, the overall F-statistics for the addition of each successive term, F and the standard deviation of regression, SD.

Table 5.2: QSAR models with the statistical parameters for 14 carbocyclic nucleosides against three cancer cell lines in gas phase.

No.	Cell line	QSAR equations	r^2	r^2_{CV}	SD	F
1a	L1210/0	$\log(IC_{50}^{-1}) = -9.102 - 2.387\ E_{NL} + 0.011\ SA$	0.962	0.941	0.131	138.92
1b		$\log(IC_{50}^{-1}) = -7.389 + 0.312\ \omega + 0.011\ SA$	0.935	0.899	0.170	80.24
2a	Molt4/C8	$\log(IC_{50}^{-1}) = -10.695 - 2.996\ E_{NL} + 0.014\ SA$	0.965	0.944	0.160	149.76
2b		$\log(IC_{50}^{-1}) = -8.520 + 0.362\ \omega + 0.014\ SA$	0.908	0.854	0.259	54.26
3a	CEM/0	$\log(IC_{50}^{-1}) = -11.119 - 3.296\ E_{NL} + 0.015\ SA$	0.958	0.934	0.184	124.40
3b		$\log(IC_{50}^{-1}) = -8.740 + 0.415\ \omega + 0.014\ SA$	0.912	0.860	0.264	57.55

A good QSAR model should have statistical parameters $r > 0.95$, $SD < 0.3$, and $r^2_{CV} > 0.60$.[24] Thus equations (1a, 2a, and 3a) with r values 0.980, 0.982, and 0.978 and equations (1b, 2b, and 3b) having r values 0.967, 0.952, and 0.955 are statistically significant. From Table 5.2, it can be seen that gas phase derived QSAR models for all three cancer cell lines explain more than 90% of the variances in the activity data along with higher values of r^2_{CV} (>0.86). However, it is observed that combinations of E_{NL} and SA values can build more significant models than that obtained by ω and SA in gas phase. Descriptors E_{NL} and ω were not considered together in the regression analysis as they are highly correlated ($r = 0.87$).

The QSAR models obtained after inclusion of solvent medium accounts for explaining 94%–96% variances of the activity data with significant values of r^2_{CV} (>0.89) (Table 5.3). In solvent phase, r^2 and r^2_{CV} values increase from that obtained in gas phase for equations 1'b, 2'b, and 3'b, where ω and SA are independent variables.

Table 5.3: QSAR models with the statistical parameters for 14 carbocyclic nucleosides against three cancer cell lines in solvent phase.

No.	Cell line	QSAR equations	r^2	r^2_{CV}	*SD*	*F*
1'a	L1210/0	$\log(IC_{50}^{-1}) = -7.734 - 2.239\ E_{NL} + 0.007\ SA$	0.949	0.921	0.151	103.19
1'b		$\log(IC_{50}^{-1}) = -6.864 + 0.355\ \omega + 0.009\ SA$	0.963	0.937	0.129	143.06
2'a	Molt4/C8	$\log(IC_{50}^{-1}) = -8.994 - 2.794\ E_{NL} + 0.010\ SA$	0.956	0.931	0.177	121.69
2'b		$\log(IC_{50}^{-1}) = -7.914 + 0.409\ \omega + 0.012\ SA$	0.937	0.894	0.213	82.69
3'a	CEM/0	$\log(IC_{50}^{-1}) = -9.263\ -3.109\ E_{NL} + 0.009\ SA$	0.947	0.914	0.205	98.95
3'b		$\log(IC_{50}^{-1}) = -8.059 + 0.470\ \omega + 0.012\ SA$	0.939	0.898	0.219	85.78

The correlation plots between experimental and calculated $\log(IC_{50}^{-1})$ values of the nucleoside analogues derived from best established QSAR models (eqs. 1a, 2a, 3a) are shown in Figure 5.2 which indicates that these descriptors can be effectively used in the prediction of cytotoxicity of carbocyclic nucleoside analogues. It is interesting to note that when ω is singularly selected, it exhibits higher positive correlation to $\log (IC_{50}^{-1})$ for L1210/0 cell line in solvent phase (r = 0.801) than in gas phase (r = 0.730). Also for other two cell lines, correlation coefficients calculated from ω have higher values in solvent medium. Thus solvent phase derived ω values can predict more reliable activities than gas phase. Similar results are obtained in Chapters 3 and 4 on *cis*-platinum complexes. Although, solvent medium does not show any influence for other equations containing E_{NL} and *SA* as independent factors, we found that solvent phase predicted E_{NL} singly can explains about 81% of variances in the activity data. We found η and E_{LUMO} (r = -0.846, -0.783, -0.798 and r = -0.67, -0.635, -0.668, respectively) are the next important parameters for the present QSAR analyses. However, the multi-linear regression analyses performed using these descriptors predict statistical parameters which are slightly less significant than that obtained in the present study.

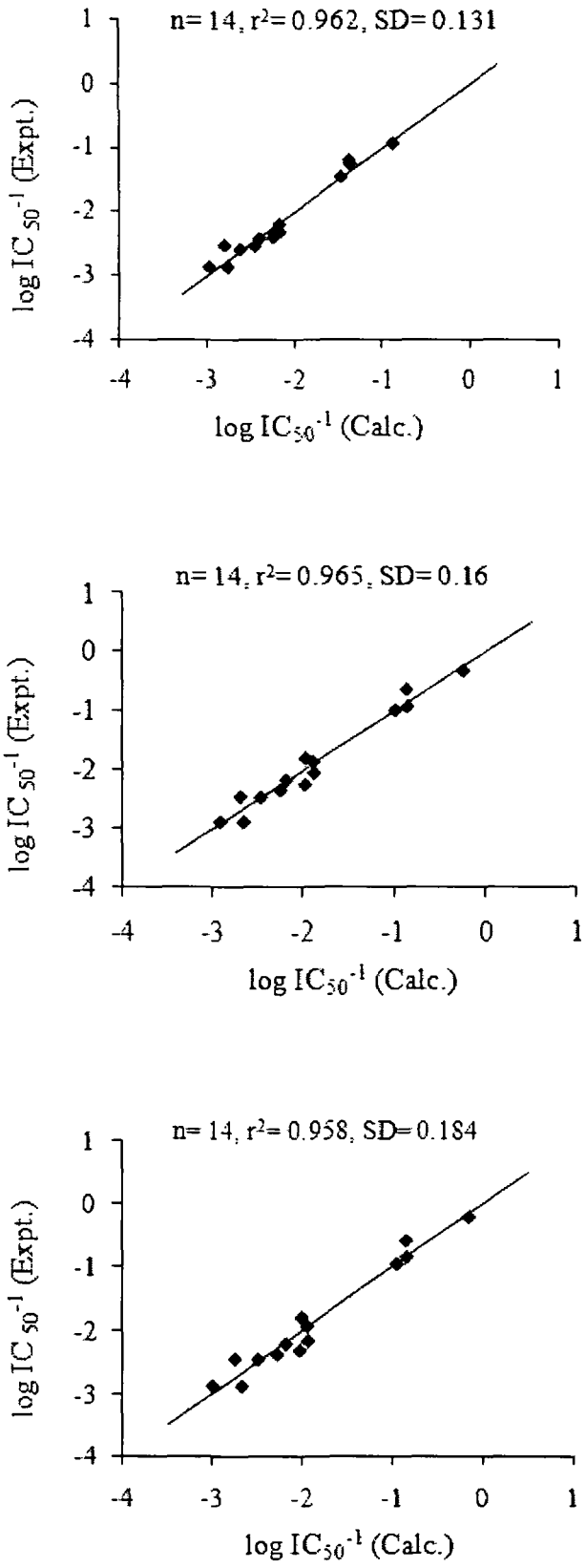


Figure 5.2: Correlation plots between experimental and calculated values of cytotoxicity ($\log(IC_{50}^{-1})$) for 14 carbocyclic nucleosides using eqs. 1a, 2a, and 3a.

In general, the anticancer drug-DNA binding mechanism involves the donation of electrons from DNA and acceptance of electrons by the drug molecule. According to the frontier molecular orbital theory[25,26], the E_{LUMO} and E_{NL} of an acceptor generally play an important role in this type of interaction by accepting electrons from the HOMO of DNA base pairs. The lower values of these parameters increase the capability of the molecules to accept electrons from the HOMO of DNA base pairs due to the orbital interaction, which results in drug–DNA association and makes the whole system stable. We found that coefficients of E_{NL} in the QSAR equations (Table 5.2 and Table 5.3) are negative, i.e, decreasing theirs values can improve the anticancer activity. From Table 5.1, we can see that in both gas and solvent phases molecules 9, 10, and 11 with lower values of E_{NL} (-1.351, -1.576, -1.311 and -1.469, -1.622, -1.477) show greater activities. Lower values of E_{NL} means more electron-withdrawing groups (e.g., halogen, -COOR, $-NO_2$ etc.) in the molecules. Chlorophenylpurine derivative (compound 10) with strong electron-withdrawing group is the most active one among the 14 molecules. Also, from Table 5.2 and Table 5.3, it is seen that the coefficients of ω are positive, so decreasing the values of E_{LUMO} will increase ω (eqs. 1.37, 1.38 and 1.39, Chapter 1) and also be beneficial to improve the activity. In addition, we observed that the coefficients of *SA* in both gas and solvent phases are positive. Thus bigger substituent on the molecules is advantageous to the improvement of their anticancer activities. For compounds 8, 9, and 11 with bigger substituent groups contribute to the increase in the surface area, thereby enhancing the activity (Table 5.1).

Based on the above discussion of the three parameters contributing to anticancer activity in QSAR equations, we can find that these parameters exhibit the great correlation to the $\log(IC_{50}^{-1})$ of cellular lines. Considering these observations, we have modified the substituent R (compounds 1-11, Figure 5.1) by large electron-windrowing group to theoretically design 10 new compounds with high anticancer activity. Table 5.4 lists their predicted activities against L1210/0, Molt4/C8, and CEM/0 cell lines from application of the QSAR models (eqs. 1a, 2a, and 3a, respectively). The $\log(IC_{50}^{-1})$ values of these 10 compounds are higher than those of the 14 carbocyclic nucleoside derivatives which indicate that our established models have strong predictive abilities and thus can be probably used in molecular design.

Table 5.4: Calculated activities for the 10 designed compounds against three cancer cell lines in gas phase.

				log(IC_{50}^{-1})		
Compd	R*	E_{NL}	*SA*	L1210/0	Molt4/C8	CEM/0
D1	C_6H_4COH	-2.054	637.45	2.812	4.383	5.212
D2	$C_6H_4COCH_3$	-1.901	417.33	0.026	0.843	1.406
D3	$C_6H_4COCH_2CH_3$	-1.873	438.65	0.194	1.057	1.634
D4	C_6H_4COOH	-1.864	645.56	2.448	3.927	4.708
D5	$C_6H_4COOCH_3$	-1.711	427.89	-0.311	0.421	0.938
D6	$C_6H_4COOCH_2CH_3$	-1.669	713.78	2.733	4.298	5.088
D7	C_6H_4COCl	-2.23	414.94	0.785	1.795	2.455
D8	$C_6H_4CF_3$	-1.738	407.25	-0.473	0.213	0.718
D9	C_6H_4CN	-1.878	397.74	-0.244	0.499	1.036
D10	$C_6H_4NO_2$	-2.639	401.71	1.616	2.835	3.604

* Substituent of compounds 1-11 in Figure 5.1

Further, we modelled the inhibitory activities of an additional set of 20 molecules (compounds 8–11 and 15–30) against murine leukemia cell line (L1210/0) with the same descriptors.[5] All the activities calibrated to the logarithmic (log (IC_{50}^{-1})) values are listed in Table 5.5 along with the used parameters.

The QSAR equations with significant values of statistical parameters in both gas and solvent phases for these 20 molecules are represented by equations 5.1 and 5.2, respectively. The values were calculated using experimental activity log(IC_{50}^{-1}) as a dependent variable and combination of three descriptors, namely ω, E_{NL}, and *SA* of the compounds with lower autocorrelation coefficients as independent variables in gas and solvent models.

$$\log(IC_{50}^{-1}) = -1.004 \ -1.381\ \omega\ -1.208\ E_{NL} + 0.009\ SA \qquad (5.1)\ \text{Gas phase}$$

$$n = 20,\ r^2 = 0.826,\ r^2_{CV} = 0.737,\ SD = 0.660,\ F = 25.33$$

$$\log(IC_{50}^{-1}) = 0.246\ -1.021\ \omega + 0.982\ E_{NL} + 0.009\ SA \qquad (5.2)\ \text{Solvent phase}$$

$$n = 20,\ r^2 = 0.846,\ r^2_{CV} = 0.693,\ SD = 0.620,\ F = 29.36$$

Table 5.5: Parameters used to build the QSAR models for 20 nucleosides analogues in gas and solvent phases.

	$\log(IC_{50}^{-1})$	Gas phase			Solvent phase		
Compd	L1210/0	ω	E_{NL}	*SA*	ω	E_{NL}	*SA*
8	-3.334	5.469	-1.410	376.77	5.851	-1.5	377.63
9	-2.734	5.271	-1.351	397.44	5.708	-1.469	397.62
10	-2.133	5.967	-1.576	393.70	5.906	-1.622	392.90
11	-2.868	5.042	-1.311	406.46	5.601	-1.477	406.46
15	-3.961	4.997	-1.552	278.19	5.112	-1.604	277.38
16	-4.559	5.664	-1.305	284.14	6.126	-1.331	283.60
17	-3.755	4.886	-1.516	327.25	5.110	-1.612	327.19
18	-0.548	2.375	-0.646	282.82	2.741	-0.625	282.88
19	-2.322	3.311	-0.645	333.92	3.915	-0.574	335.51
20	0.562	2.693	-1.103	276.80	2.290	-0.417	277.20
21	-2.008	3.016	-0.843	269.21	3.656	-1.084	267.72
22	-1.782	3.060	-0.963	257.29	3.458	-1.186	255.65
23	-0.182	2.041	-1.331	257.34	2.450	-0.728	255.85
24	-3.989	5.886	-1.734	407.56	5.305	-1.944	406.12
25	-4.644	5.845	-1.503	290.83	5.659	-1.486	286.68
26	-4.963	6.090	-1.505	283.29	5.470	-1.385	282.91
27	-3.761	5.071	-1.664	298.33	5.390	-1.321	295.74
28	-3.219	4.762	-1.654	282.68	4.689	-1.310	282.10
29	-3.600	5.525	-2.152	246.46	5.033	-1.301	246.16
30	-3.282	5.203	-1.752	244.95	5.060	-1.295	244.73

The plots between experimental and calculated values of $\log(IC_{50}^{-1})$ predicted by gas and solvent phases are presented in Figure 5.3. These plots suggest that the selected descriptors can be effectively used in determination of activities of nucleoside analogues.

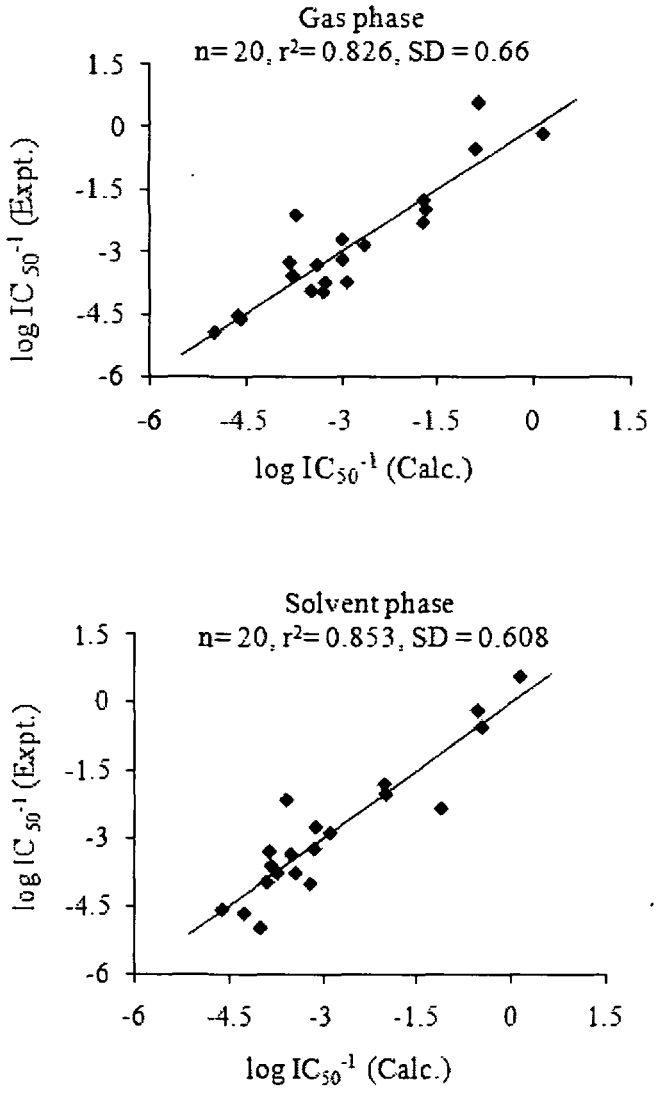


Figure 5.3: Plots between experimental versus calculated values of cytotoxicity ($\log(IC_{50}^{-1})$) for 20 nucleoside analogues in both gas and solvent phases

5.5 Conclusions

The 2D QSAR studies only with two parameters performed for 14 carbocyclic nucleosides against three cancer cell lines in gas and solvent phases show good statistical quality both in regression ($r^2 > 0.90$) and LOO cross-validation

($r^2_{CV} > 0.86$). The regression models reveal that lower values of E_{NL} combined with higher values of ω and SA, increase inhibitory activities against three cancer cell lines. Based on the key features of the molecules that are necessary for their anticancer activity, 10 new compounds with rather high anticancer activities against L1210/0, Molt4/C8, and CEM/0 cell lines than those of 14 compounds have been theoretically designed. The QSAR models developed for an additional set of 20 nucleoside analogues with three parameters i.e., ω, E_{NL}, and SA provide significant statistical parameters in both gas and solvent media. Moreover, the presented QSAR models have number of variables which is seven times less than the number of observations. Hence, the current work clearly shows the effectiveness of these DFT and MM+ derived parameters in QSAR analysis of nucleoside analogues.

References

(1) Jordheim, L.; Galmarini, C.M.; Dumontet, C. Nucleoside analogues and nucleobases in cancer treatment. *Lancet Oncol* **3**, 415–424 (2002)

(2) Jordheim, L.; Galmarini, C.M.; Dumontet, C. Recent developments to improve the efficacy of cytotoxic nucleoside analogues. *Anti-Cancer Drug Discovery* **1**, 163–170 (2006)

(3) Cheson, B.D.; Keating, M.J.; Plunkett, W. *Nucleoside Analogues in Cancer Therapy* (Marcel Dekker, New York, 1997)

(4) David, Y.B. *et al* Complementary antineoplastic activity of the cytosine nucleoside analogues troxacitabine (Troxatyl) and cytarabine in human leukemia cells. *Cancer Chemother. Pharmacol.* **52**, 497–506 (2003)

(5) Helguera, A.M.; Rodriguez-Borges, J.E.; Garcia-Mera, X.; Fernandez, F.; Cordeiroa, M.N.D.S. Probing the anticancer activity of nucleoside analogues: A QSAR model approach using an internally consistent training set. *J. Med. Chem.* **50**, 1537–1545 (2007)

(6) Bressi, J.C. *et al* Adenosine analogues as inhibitors of *Trypanosoma brucei* phosphoglycerate kinase: Elucidation of a novel binding mode for a 2-Amino-N6-substituted adenosine. *J. Med. Chem.* **43**, 4135–4150 (2000)

(7) Vince, R.; Hua, M. Synthesis and anti-HIV activity of carbocyclic 2',3'-didehydro-2',3'-dideoxy 2,6-disubstituted purine nucleosides. *J. Med. Chem.* **33**, 17–21 (1999)

(8) Nieto, M.I. *et al* Synthesis, antiviral and cytostatic activities of carbocyclic nucleosides incorporating a modified cyclopentane ring. Part 2: adenosine and uridine analogues. *Nucleosides Nucleotides* **17**, 1255–1266 (1998)

(9) Figueira, M.J. *et al* Synthesis and antiviral and cytostatic activities of carbocyclic nucleosides incorporating a modified cyclobutane ring. Part 1. Guanosine analogues. *Arch. Pharm.* **332**, 348–352 (1999)

(10) López, C. *et al* A cyclobutane carbonucleoside with marked selectivity against TK+ and TK- varicella zoster virus. *Nucleosides Nucleotides Nucleic Acids* **20**, 1133–1135 (2001)

(11) Figueira, M.J.; Caamaño, O.; Fernández, F.; Blanco, J.M. Synthesis of (±)-*c*-4-amino-*r*-1,*t*-2,*c*-3-cyclopentanetrimethanol and higher homologues of 8-

azapurine arabino-carbocyclic nucleosides. *Tetrahedron*, **58**, 7233–7240 (2002)

(12) Yao S.W. *et al* Synthesis and QSAR study of the anticancer activity of some novel indane carbocyclic nucleosides. *Bioorg. Med. Chem.* **11**, 4999–5006 (2003)

(13) Caamaño, O. *et al* A Convenient synthesis of new purinylhomo-carbonucleosides on a cyclopentane ring fused with pyridazine. *Synthesis* **17**, 2855–2862 (2004)

(14) Fernández, F.; García-Mera, X.; López, C.; Morales, M.; Rodríguez- Borges, J.E. A Convenient synthesis of novel pyrimidinyl-5′-nor-1′ homocarbanucleosides based on indanol. *Synthesis* **20**, 3549–3554 (2005)

(15) Fernández, F.; García-Mera, X.; Morales, M.; Rodríguez- Borges, J.E. A short, efficient synthesis of substituted uracil: An indane carbocyclic nucleoside *Synthesis* **2**, 239–242 (2001)

(16) Abad, F.; Alvarez, F.; Fernández, F.; García-Mera, X.; Rodríguez- Borges, J.E. New carbocyclic nucleosides derived from indan. *Nucleosides Nucleotides Nucleic Acids* **20**, 1127–1128 (2001)

(17) Santana, L.; Teijeira, M.; Uriarte, E.; Balzarini, J.; De Clercq, E. Synthesis, conformational analysis and antiviral and antitumoral activity of new 1,2-disubstituted carbocyclic nucleosides. *Eur. J. Med. Chem.* **37**, 755–760 (2002)

(18) Huuskonen, J. QSAR modeling with the electrotopological state: TIBO derivatives. *J. Chem. Inform. Comput. Sci.* **41**, 425–429 (2001)

(19) Chen, J.C.; Qian, L.; Wu, W.J.; Chen, L.M.; Zheng, K.C. A QSAR study of substituted benzo[*a*]phenazines as potential anticancer agents. *J. Mol. Struct.: THEOCHEM* **756**, 167–172 (2005)

(20) Chen, J.; Shen, Y.; Liao, S.; Chen, L.; Zheng, K. DFT-based QSAR study and molecular design of AHMA derivatives as potent anticancer agents. *Int. J. Quant. Chem.* **107**, 1468–1478 (2007)

(21) Wan, J.; Zhang, L.; Yang, G.F. Quantitative structure-activity relationship for phenyl triazolinones of protoporphyrinogen oxidase inhibitors: A density functional theory study. *J. Comput. Chem.* **25**, 1827–1832 (2004)

(22) Srivastava, H.K.; Pasha, F.A.; Singh, P.P. Atomic softness-based QSAR study of testosterone. *Int. J. Quant. Chem.* **103**, 237–245 (2005)

(23) Karelson, M.; Lobanov, V.S. Quantum-chemical descriptors in QSAR/QSPR studies. *Chem. Rev.* **96**, 1027–1043 (1996)

(24) Wold, S. Validation of QSAR's. *Quant. Struct. Act. Rel.* **10**, 191–193 (1991)

(25) Fukui, K.; Yonezawa, T.; Shingu, H. A molecular orbital theory of reactivity in aromatic hydrocarbons. *J. Chem. Phys.* **20**, 722–725 (1952)

(26) Fleming, J. *Frontier Orbital and Organic Chemical Reaction* (Wiley, New York, 1976)

Chapter 6

Hydrolysis of AMD473 (*cis*-[$PtCl_2(NH_3)$(2-picoline)]) and Its Interactions with Guanine

In this chapter we have investigated the hydrolysis of cisplatin analogue *cis*-[$PtCl_2(NH_3)$(2-picoline)] (AMD473) using HF and DFT calculations. Four different paths of hydrolysis have been studied by considering the replacement of two Cl atoms *trans* to NH_3 and 2-picoline ligands. The geometries of different reactant and product complexes are confirmed by IRC calculations from the transition state structures for each reaction. The results indicate that the rate of hydrolysis of chloride ligand *trans* to 2-picoline group is higher than that of chloride ligand *cis* to 2-picoline group due to steric effect experienced by the axial picoline ligand. The rate constants calculated in gas phase for the first steps of hydrolysis reactions are in close agreement with the experimental values. However, the gas phase values for the second steps differ significantly from the experimental data. An improvement of these values is made by incorporating solvent medium in our calculations. Further, we have performed monofunctional binding of the obtained aquo species with guanine to provide a detailed understanding of binding mechanism of this anticancer drug. The HF/6-31G* calculated rate constants, $k_3 = 1.43\times10^{-3}$ s^{-1} M^{-1} and $k_{3'} = 5.42\times10^{-3}$ s^{-1} M^{-1} are in good agreement with the experimental values, 6.67×10^{-3} s^{-1} M^{-1} and 7.97×10^{-3} s^{-1} M^{-1}, respectively. [Sarmah and Deka, *J. Mol. Struct. THEOCHEM* **955**, 53–60 (2010)]

6.1 Introduction

The successful development of platinum complexes for their effect as anticancer drugs began with *cis*-Diamminedichloroplatinum(II), clinically known as cisplatin,[1] the anticancer activity of which was discovered by Rosenberg and co-workers.[2] Although the great success in treating certain kinds of cancer, such as testicular and ovarian cancers,[3,4] cisplatin does have some limitations. The critical drawbacks of cisplatin as an anticancer drug include its toxic side effects, intrinsic and acquired cellular resistance and limited solubility in aqueous solution. Over the years, various platinum complexes have been studied in an attempt to overcome these problems. Unfortunately, the second generation drugs, carboplatin and oxaliplatin, also suffer from drug resistance and other side effects.[5,6] Platinum(II) complexes with bulky planar ligands, such as pyridine and substituted pyridine have shown to reduce the rate of deactivation by sulfhydryl groups without affecting the cytotoxic activity.[7] One such promising anticancer agent, *cis*-$[PtCl_2(NH_3)$(2-picoline)], known as AMD473 has now entered the worldwide phase II and III clinical trials.[8] This orally administrated drug is less toxic than cisplatin and possesses activity against cisplatin-resistant cell lines.

It is generally accepted that platinum (II) drugs bind with DNA, forming most abundantly 1,2-intrastrand cross-links between the N7 atoms of two adjacent guanine units.[3] However, before reaching DNA, these molecules undergo hydrolysis within the cell, where chloride concentration ranges between 2 and 30 mM. In this range, one or both the chloride leaving groups are substituted by water molecules forming monoaquo or diaquo products, respectively. These aquo species are more reactive than their parent compounds towards nucleophilic centers of bio-molecules. AMD473 undergoes hydrolysis [eq. 6.1 and eq. 6.2] very slowly compared to that of cisplatin due to the steric hindrance experienced by 2-methyl group of picoline ligand of the molecule.

$$[PtCl_2(NH_3)(2\text{-pic})] + H_2O \rightarrow [PtCl(H_2O)(NH_3)(2\text{-pic})]^+ + Cl^- \quad (6.1)$$

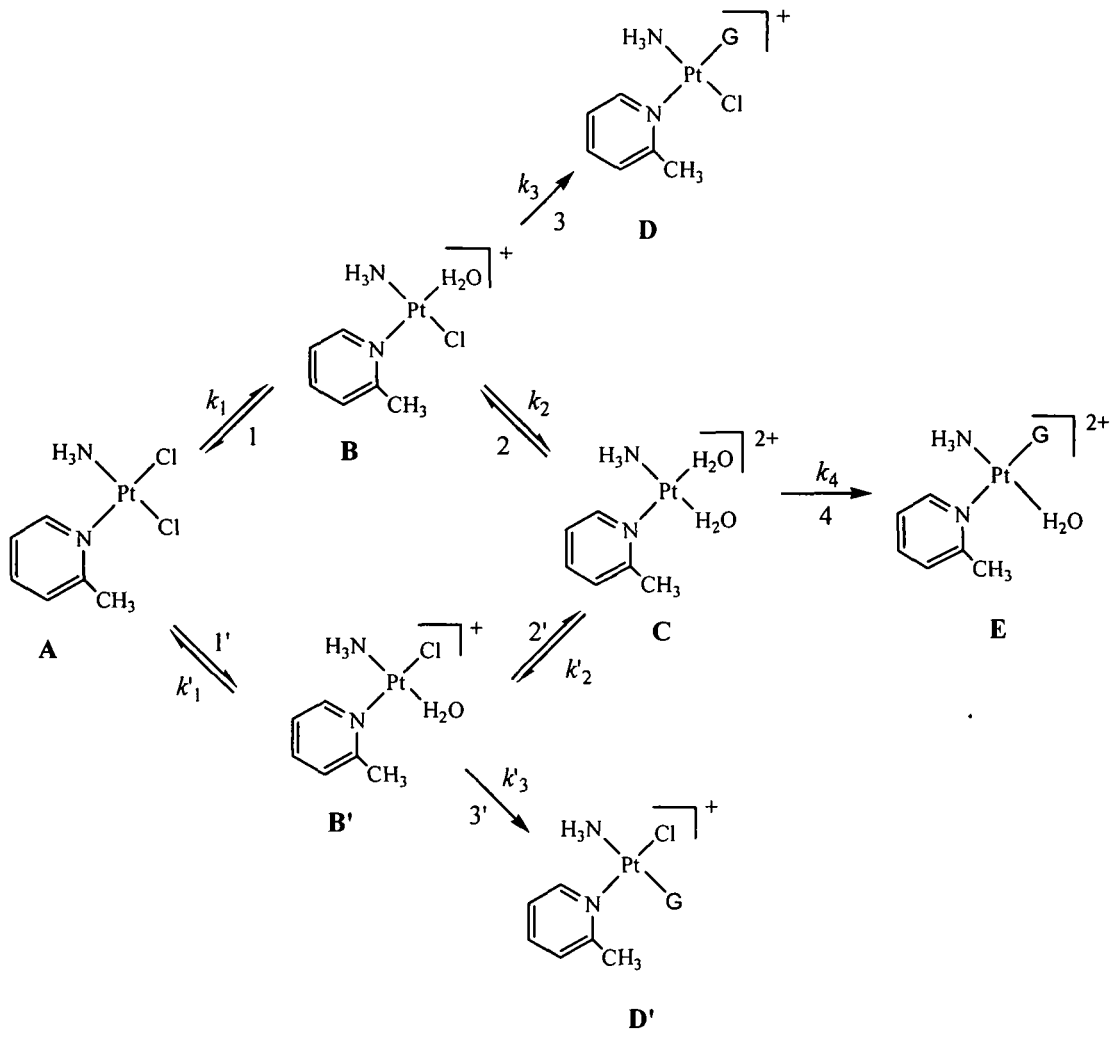
$$[PtCl(H_2O)(NH_3)(2\text{-pic})]^+ + H_2O \rightarrow [Pt(H_2O)_2(NH_3)(2\text{-pic})]^{2+} + Cl^- \quad (6.2)$$

Sadler *et al.*[9] studied hydrolysis of this drug molecule along with its 3-picoline analogue using [^{1}H, ^{15}N] 2D NMR spectroscopy and found that the first-step

hydrolysis rates for both the complexes are slower than that of cisplatin. The final target of monoaquo and diaquo products of platinum drug is DNA. However, on their way to the target, these species can also interact with S-containing biomolecules such as cysteine and methionine residues causing drug inactivation.

Over the last few years, many theoretical studies have been carried out on cisplatin-DNA interaction to provide detailed insight into the binding mechanism at molecular level. The structural and spectral analyses of cisplatin and some of its analogues were investigated by Wysokinski and Michalska using different density functional models.[10,11] The substitution reactions of chloride ligands of cisplatin by water molecules have been performed at different theoretical levels, which suggest trigonal bipyramidal like structure for the transition states of these reactions.[12-18] The kinetic and thermodynamic analyses for binding of reactive aquo complexes of cisplatin with guanine and adenine provides detailed reaction energy profile as well as structure of the transition states for these ligand exchange reactions.[19-21] Burda and his group studied the geometry of cisplatin substituted by one or two purine bases using DFT calculations[22] and reported that sugar-phosphate chain onto Pt-bridged fragments enhances the stability of these complexes.[23] Matsui *et al.*[24,25] examined the changes of hydrogen bonds between two or more base pairs due to binding of cisplatin through proton transfer reactions which causes mispairing of the pairs. Interactions of cisplatin with DNA GpG sequence in water were investigated using Car-Parinello MD calculations.[26] Further, kinetic factors governing the affinity of cisplatin and its dinuclear analogues towards nitrogen and sulfur nucleophiles have been studied by Deubel.[27,28]

The objective of present study is to explore the detailed mechanistic behaviour of sterically hindered anticancer drug, *cis*-[$PtCl_2(NH_3)$(2-picoline)] (A) with different theoretical approach, focusing the reactions shown in Scheme 6.1. The reactant drug molecule A undergoes first hydrolysis *via* step 1 and step 1' to give products B and B' that further undergo second hydrolysis through steps 2 and 2' and produce diaquated form C. Binding of guanine with two monoaquated forms of A proceeds *via* step 3 and step 3' to give products D and D'. The substitution reaction of diaquated form with guanine is also studied where the product E is obtained.



Scheme 6.1: Reaction scheme for hydrolysis and guanine binding steps of AMD473 (A).

6.2 Computational Details

The geometries of molecules and transition states involved in the hydrolysis process were optimized using HF and DFT with four different hybrid GGA exchange correlation functionals: B3PW91, B3P86, mPW1PW91 and PBE1PBE in Gaussian 03 program.[29] By noting that our HF results for hydrolysis reactions were comparable with the available experimental data and overall best performance of this method on cisplatin like molecules,[30] we performed three ligand exchange reactions of AMD473 with guanine (steps 3, step3', and step 4, Scheme 6.1) at this level. B3PW91 and B3P86 are Becke's three-parameter functionals where the nonlocal correlation is provided by Perdew-Wang91[31] and Perdew's 86,[32] respectively. The

mPW1PW91 is the combination of modified PW exchange and PW91 correlation[33] functionals suggested by Adamo and Barone.[34] PBE1PBE is the exchange-correlation functionals of Perdew, Burke, and Ernzerhof (PBE).[35] We used double-zeta LanL2DZ ECP basis set as described by Hay and Wadt[36] for Pt atom which incorporates the massvelocity and Darwin relativistic effects into the potential. All other atoms of the complexes were treated with standard split valance basis set 6-31G*.[37] The nature of the stationary points located on the potential energy surface (PES) was checked by vibrational analysis, i.e., all positive frequencies ensuring minimal state and one negative frequency ensuring first-order transition state (TS). For all TS, intrinsic reaction coordinate (IRC) calculations were performed according to Gonzalez and Schlegel[38,39]scheme as implemented in the used program and the species connected by the IRC were fully optimized to get the respective intermediate structures. Thermal contributions to each Gibbs free energy were considered at 298.15 K and 1 atm.

The rate constants (k) for all reactions were calculated using eq. 6.3 according to the Eyring's transition state theory (TST). In this equation, k_B, T, and h are the Boltzmann constant, absolute temperature, and Planck constant, respectively. $\Delta G^{\ddagger}$ is the activation Gibbs free energy for each reaction.

$$k(T) = \frac{k_B T}{h} e^{-\Delta G^{\ddagger} / RT} \quad (6.3)$$

The influence of solvent phase on the energetics of the reactions was investigated using the polarizable continuum model, IEF–PCM.[40-42] Solvent effects were accounted by means of single-point calculations of the complexes in solution as this approach was found to be good for such systems.[21,43] A dielectric constant of 78.39 for water was used to approximate the bulk effects of solvation.

6.3 Results and Discussion

The calculated structural parameters of reactants, transition states, and products obtained during hydrolysis of AMD473 are quite similar at all levels of theoretical methods used in the present study. However, kinetic results of HF/6-31G* and mPW1PW91/6-31G* methods are more comparable with the experimental data[9] than the other levels of calculations. Therefore, all the results of hydrolysis reactions are discussed at mPW1PW91/6-31G* level.

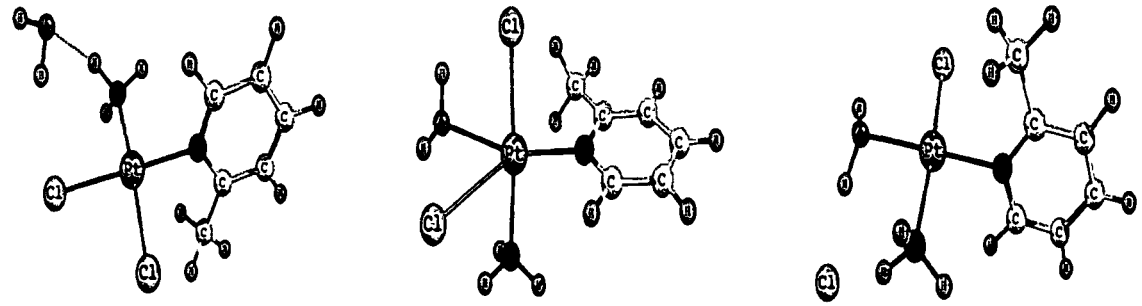
The optimized geometry of the drug molecule (A, Scheme 6.1) has a square-planar structure in agreement with X-ray crystal structure reported by Chen *et al.*[9] The Pt–Cl bond (2.33Å) *trans* to the picoline group is longer than the Pt–Cl bond (2.32Å) in *cis* position making the former bond weaker compared to the later one. This observation is in agreement with the results reported by Michalska and Wysokinski.[11] The Pt–N bond lengths for NH_3 and 2-picoline groups are 2.06Å and 2.03Å, respectively. The picoline ring is found to be perpendicular to the molecular plane with H_3C–N(picoline)–Pt–N(ammonia) dihedral angle equals to 102.3° and H_3C– Pt distance equals to 3.19 Å in agreement with previous theoretical and experimental results.[44,9]

6.3.1 Hydrolysis

6.3.1.1 Geometric Profiles

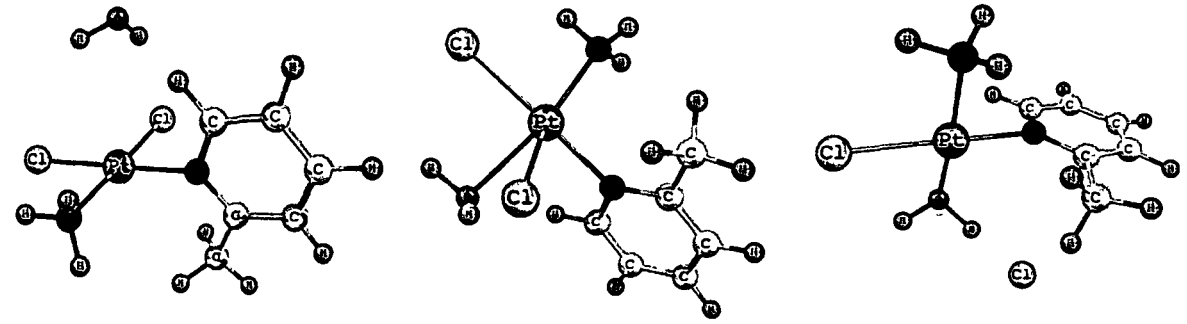
The mPW1PW91 level predicted stationary points along the reaction coordinates of first hydrolysis reactions are presented in Figure 6.1. The reactant complex (RC) for the first hydrolysis of chloride ligand *trans* to picoline (step 1), RC-1, possessed water molecule between the amine group and chlorine atom through hydrogen bonds which connect the hydrogen of water with chloride atom ($r_{HOH\cdots Cl}$= 2.30Å) and oxygen of water with an amine hydrogen ($r_{H2O\cdots HNH2}$ = 1.80Å). In contrast to RC-1, the water molecule approaches the reaction centre in reactant complex of step 1' (RC-1') *via* hydrogen bonds with a hydrogen atom of picoline ring (2.21Å) and chlorine atom *trans* to the amine ligand (2.43Å). Comparison of absolute energies for the two structures reveals that RC-1 is 7.1 kcal/mol more stable than RC-1'. The transition state (TS) structures for the first step hydrolysis of chloride ligand *trans* to picoline (TS-1) and *trans* to amine (TS-1'), have distorted trigonal-bipyramidal arrangement as expected for such platinum complexes involving associative mechanism. The bond lengths for the bond being broken between Pt and Cl and the newly formed bond between Pt and O of entering water molecule are 2.73 Å and 2.44 Å, respectively. However, in TS-1', the Pt–Cl distance is slightly longer and Pt–O bond is shorter than that of TS-1. In both the TS structures, bond angles between entering and leaving ligands (68.9° and 67.8°) are in good agreement with the value reported for cisplatin.[13]

Step 1



RC-1 TS-1 PC-1

Step 1'



RC-1' TS-1' PC-1'

Figure 6.1: Stationary points (RC, TS, and PC) along the reaction coordinate of first hydrolysis reactions optimized at mPW1PW91/6-31G* level of theory.

The details geometrical parameters of TS-1 and TS-1' calculated at each level of theory are provided in the Table 6.1. It is interesting to note that TS-1 is more hydrogen bonded than TS-1' with two hydrogen bonds between chlorine atom and hydrogen atoms of aqua (2.13Å) and amine (2.31Å) groups which accounts for the greater stability of the structure (2.9 kcal/mol). The orientation of picoline ring changes significantly in TS-1' making tilt of the ring smaller as indicated by the H_3C–N(picoline)–Pt–N(H_3) angle of about 71.2°. The intermolecular interaction of the chlorine atom in both the product complexes (PC) differs significantly. In PC-1, it is hydrogen bonded with aqua ligand and amine group quite similar to the intermediate structure calculated for cisplatin.[12] An internal rotation around the Pt–N(picoline) bond gives PC-1' where chloride leaving group exhibits interaction with one hydrogen atom of the methyl group as well as with aqua ligand. For both TS-1

and TS-1', the corresponding structures of reactants and products are confirmed by IRC calculations.

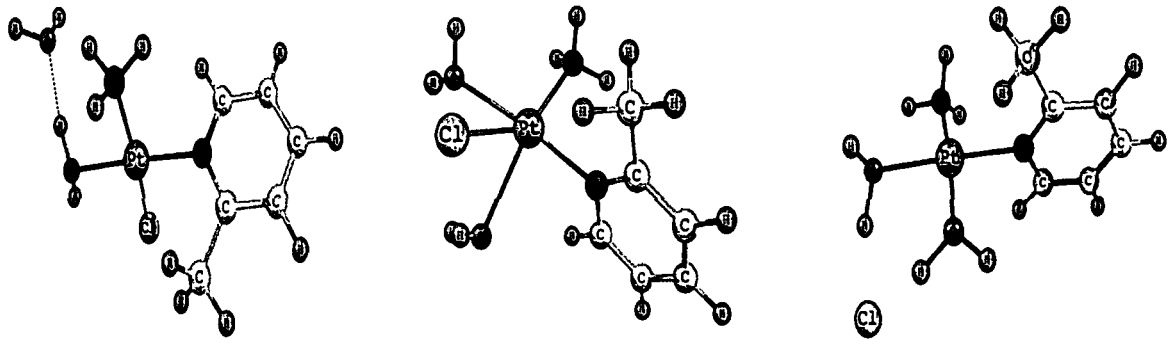
Table 6.1: Structural parameters for the transition states (TS) in the first hydrolysis of AMD473. (The unit of distance is angstroms, the unit of angle is degrees).

	HF		B3PW91		B3P86		mPW1PW91		PBE1PBE	
	TS-1	TS-1'	TS-1	TS-1'	TS-1	TS-1'	TS-1	TS-1'	TS-1	TS-1'
Pt–O	2.501	2.409	2.453	2.355	2.443	2.347	2.441	2.346	2.438	2.345
Pt–Clax	2.370	2.370	2.340	2.345	2.336	2.341	2.335	2.339	2.333	2.337
Pt–Cleq	2.790	2.850	2.748	2.791	2.734	2.776	2.733	2.778	2.728	2.772
Pt–Nax	2.101	2.091	2.078	2.045	2.072	2.039	2.071	2.041	2.069	2.038
Pt–Neq	2.081	2.101	2.024	2.074	2.024	2.069	2.024	2.067	2.021	2.063
Clax–Pt–O	85.7	87.1	84.1	86.7	83.8	86.8	84.2	86.8	83.9	86.7
Cleq–Pt–O	69.9	68.6	68.7	67.6	68.7	67.7	68.9	67.8	68.9	68.0
Nax–Pt–O	90.9	90.8	91.9	90.5	92.1	90.3	92.0	90.4	92.2	90.4
Neq–Pt–O	151.3	152.1	151.1	152.9	150.7	152.6	151.0	152.8	150.9	152.7

Second hydrolysis proceeds *via* step 2 and step 2' through interaction of water molecule with the mono-substituted aqua complex (Figure 6.2). RC-2 exhibits two strong hydrogen bonds (2.11Å and 1.71Å) resulting from the interaction of water molecule with amine ligand and aqua ligand. The structure of RC-2' differs from RC-2 where water molecule coordinates with NH_3 group (1.72Å) along with the chloride ligand through a weak hydrogen bond (2.57Å). However, RC-2 is only about 1.9 kcal/mol more stable than RC-2'. We also take into account the possibilities of forming reactant complexes for second hydrolysis by considering the water molecule between chloride and aqua ligands. These structures are found to be more stable than RC-2 and RC-2'. However, IRC calculations correspond to the structures presented in Figure 6.2 and all properties calculated with these structures are more comparable. Second hydrolysis step also proceeds through a distorted trigonal bipyramidal transition state structure. The bond angles between entering

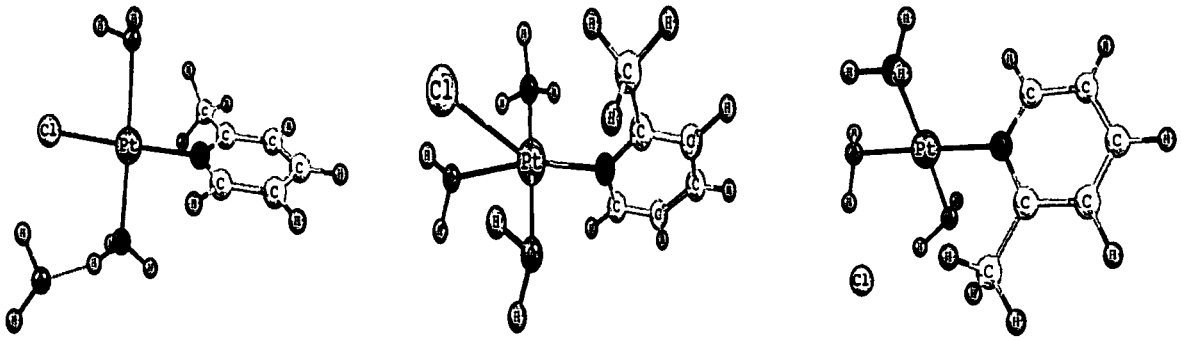
ligand, platinum atom, and leaving ligand for both transition state structures of second hydrolysis (TS-2 and TS-2') are in good agreement with the expected value (~70°).[13]

Step 2



RC-2 TS-2 PC-2

Step 2'



RC-2' TS-2' PC-2'

Figure 6.2: Stationary points (RC, TS, and PC) along the reaction coordinate of second hydrolysis reactions optimized at mPW1PW91/6-31G* level of theory.

Further details of geometrical parameters of the structures are listed in the Table 6.2. The product complexes PC-2 and PC-2' are diaquo forms of the drug having hydrogen bonded chlorine atom. Both the structures are confirmed by IRC calculations. For all levels of theory we found the product obtained by first hydrolysis of Cl^- *trans* to picoline group (step 1) is more stable.

Table 6.2: Structural parameters for the transition states (TS) in the second hydrolysis of AMD473. (The unit of distance is angstroms, the unit of angle is degrees).

	HF		B3PW91		B3P86		mPW1PW91		PBE1PBE	
	TS-2	TS-2'	TS-2	TS-2'	TS-2	TS-2'	TS-2	TS-2'	TS-2	TS-2'
Pt– Oax	2.125	2.119	2.121	2.100	2.117	2.096	2.111	2.097	2.116	2.096
Pt–Oeq	2.387	2.368	2.034	2.390	2.332	2.380	2.370	2.381	2.330	2.380
Pt–Cl	2.790	2.784	2.750	2.770	2.740	2.751	2.720	2.750	2.730	2.731
Pt–Nax	2.045	2.041	2.000	2.031	2.010	2.020	1.999	2.020	2.001	2.021
Pt–Neq	2.110	2.060	2.080	2.020	2.081	2.021	2.060	2.020	2.081	2.010
Oax–Pt–Cl	73.9	76.1	70.7	71.2	70.9	71.2	72.0	71.7	71.2	71.8
Oeq–Pt–Cl	70.4	69.8	69.3	67.8	69.4	67.9	69.4	68.4	69.7	68.0
Nax–Pt–Cl	104.5	97.4	104.6	103.4	104.2	103.1	106.0	102.9	104.0	102.8
Neq–Pt–Cl	131.9	131.7	138.7	137.1	138.8	137.9	136.1	136.6	138.4	136.7

6.3.1.2 Kinetic Analysis

Change of Gibbs free energies for all steps of hydrolysis reactions obtained at different theoretical methods are presented in Table 6.3. Our gas phase Gibbs energy (22.9 kcal/mol) is in good agreement with the experimental value.[9] Incorporation of solvent phase by using PCM model lowers the Gibbs free energies for all the steps (Table 6.3). It is interesting to note that replacement of first chloride anion *trans* to picoline group by water *via* step 1 proceeds with lower activation barriers at all studied levels of theory than that of step 1' (Cl^- *trans* to amine). This may be expected due to higher trans influence of NH_3 versus 2-picoline. The gas phase predicted rate constants for first hydrolysis reactions (Table 6.3) are in good agreement with the experimental values[9] compared to that found in aqueous medium.

Table 6.3: Gibbs free energies and rate constants for the hydrolysis reactions of AMD473. $\Delta G^{\ddagger}$ values are in kcal mol^{-1}. Solvent phase values are in parenthesis.

	HF	B3PW91	B3P86	mPW1PW91	PBE1PBE	Expt.[a]
First steps						
$\Delta G_1^{\ddagger}$	23.15	22.80	22.59	22.94	23.09	
	(20.46)	(19.28)	(20.58)	(20.22)	(20.79)	
$\Delta G_{-1}^{\ddagger}$	16.63	18.70	19.18	18.71	18.84	
$\Delta G_{1'}^{\ddagger}$	24.26	22.99	22.69	23.12	23.11	
	(21.25)	(20.03)	(21.35)	(20.83)	(21.12)	
$\Delta G_{-1'}^{\ddagger}$	13.39	14.83	14.90	14.12	14.38	
k_1	6.64×10^{-5}	1.19×10^{-4}	1.69×10^{-4}	9.39×10^{-5}	7.23×10^{-5}	3.19×10^{-5}
	(6.22×10^{-3})	(4.56×10^{-2})	(5.08×10^{-3})	(9.34×10^{-3})	(3.56×10^{-3})	
$k_{1'}$	1.01×10^{-5}	8.68×10^{-5}	1.42×10^{-4}	6.98×10^{-5}	7.01×10^{-5}	2.21×10^{-5}
	(1.64×10^{-3})	(1.28×10^{-2})	(1.38×10^{-3})	(3.33×10^{-3})	(2.04×10^{-3})	
Second step						
$\Delta G_2^{\ddagger}$	32.61	32.23	31.09	31.99	33.26	
	(21.56)	(24.64)	(25.22)	(24.85)	(24.27)	
$\Delta G_{-2}^{\ddagger}$	5.32	13.38	13.25	12.86	14.63	
$\Delta G_{2'}^{\ddagger}$	30.58	29.35	28.72	29.38	29.01	
	(22.48)	(20.46)	(21.42)	(26.52)	(21.77)	
$\Delta G_{-2'}^{\ddagger}$	6.47	12.93	13.37	12.75	12.87	
k_2	7.63×10^{-12}	1.44×10^{-11}	1.00×10^{-10}	2.19×10^{-11}	2.54×10^{-12}	7.30×10^{-5}
	(9.52×10^{-4})	(5.28×10^{-6})	(1.98×10^{-6})	(3.72×10^{-6})	(9.94×10^{-6})	
$k_{2'}$	2.34×10^{-10}	1.86×10^{-9}	5.43×10^{-9}	1.78×10^{-9}	3.31×10^{-9}	3.50×10^{-6}
	(2.02×10^{-4})	(6.16×10^{-3})	(1.21×10^{-3})	(2.21×10^{-7})	(6.79×10^{-4})	

[a] From Ref. 9

Figure 6.3 shows the reaction pathways of first (Figure 6.3. a) and second hydrolysis (Figure 6.3. b) reactions calculated at gas phase mPW1PW91/6-31G* level.

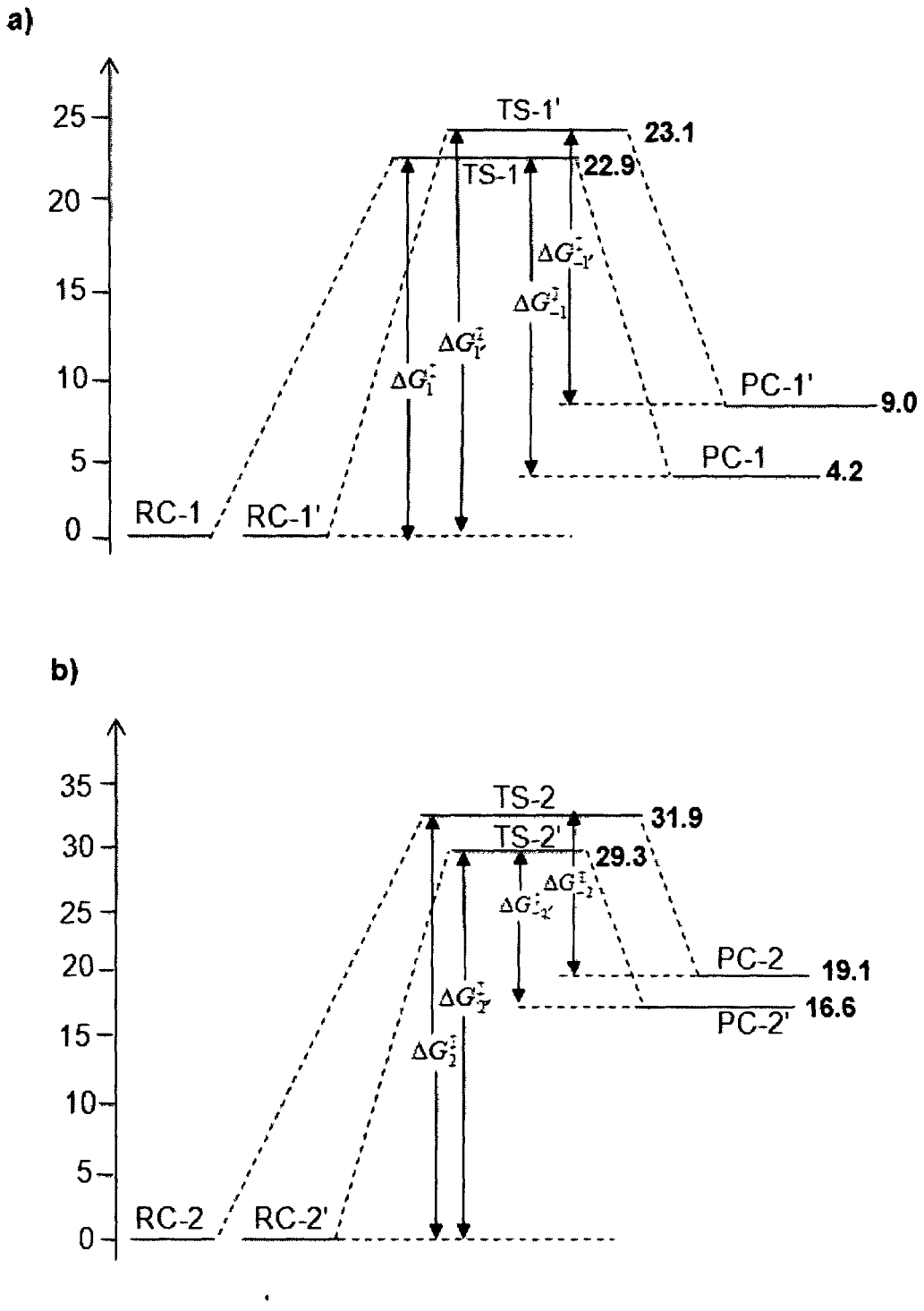


Figure 6.3: Potential energy profile of the first (a) and second (b) hydrolysis reactions calculated at mPW1PW91/6-31G* level of theory.

The first hydrolysis step is endothermic by 4.2 kcal/mol and 9 kcal/mol for step 1 and step 1', respectively. The PC-1 lies about 4.7 kcal/mol lowers in energy than PC-1'. The activation barrier for step 2 is 24.8 kcal/mol in water (31.9 kcal/mol for

gas phase), and that for step 2' is 26.5 kcal/mol (29.3 kcal/mol in vacuo, Table 6.3). Similar to the first step, the second step is also endothermic (Figure 6.3. b). We found that the gas phase predicted rate constants for the second steps of hydrolysis differ significantly from their experimental values. However, these values calculated with solvent model (Table 6.3) are in reasonable agreement with experimental values.[9] This observation is in agreement with the study reported by Costa *et al.*[14] where the rate constant for the second hydrolysis of $Pt(en)Cl_2$ in aqueous medium (SCRF-HF) was in accordance with the experimental value than that of their gas phase values.

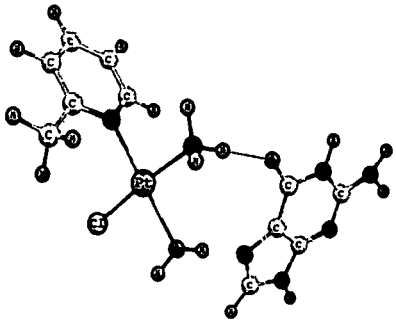
These results predict different hydrated form of *cis*-$[PtCl_2(NH_3)(2\text{-pic})]$ due to its asymmetric structure which may give rise to several isomers when bind with DNA. Similar to cisplatin, the second step is slower than the first step and thus the drug should reach DNA in its monohydrated form. This observation is in agreement with the experimental report by Chen *et al.*[45] where they found that the drug AMD473 undergoes initial hydrolysis followed by guanine substitution to give monofunctional adduct and then further hydrolysis and guanine binding to give the final bifunctional adduct. Between the two paths studied for first hydrolysis, step 1 is found energetically and kinetically more favorable than step 1'. Further monofunctional binding of these hydrated species with guanine is studied to give detailed insight into binding mechanism of this sterically hindered drug.

6.3.2 Binding with Guanine

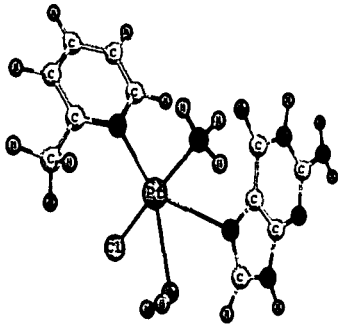
6.3.2.1 Geometric Profiles

The HF/6-31G* optimized stationary points along the reaction coordinates of all ligand exchange reactions of AMD473 with guanine are presented in Figure 6.4. Stabilization of RC-3 arises due to the presence of hydrogen bonding between NH_3 ligand and O6 atom of guanine (1.91Å), and H_2O ligand and N7 atom of guanine (1.82Å). The TS-3 structure having distorted trigonal-bypiramidal geometry is characterized as a first-order transition state. The bond length for the entering guanine molecule in this structure is 2.68Å, which is longer than that for the leaving water molecule (2.52Å).

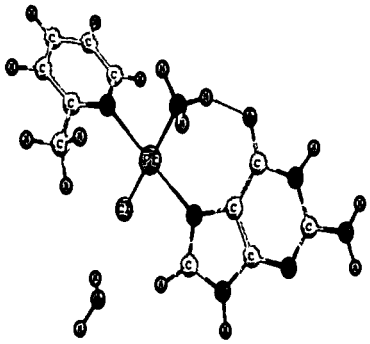
Step 3



RC-3

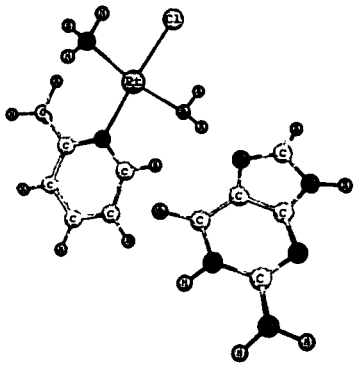


TS-3

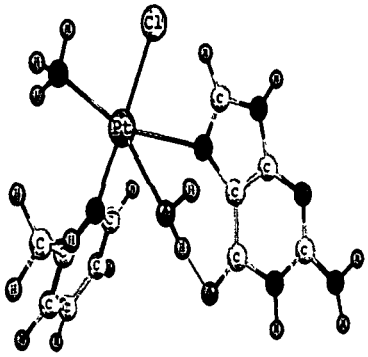


PC-3

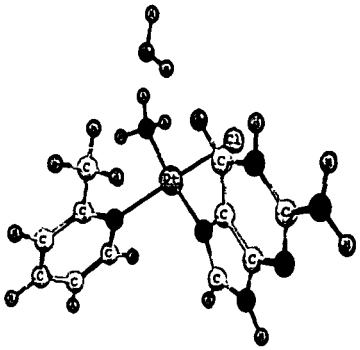
Step 3'



RC-3'

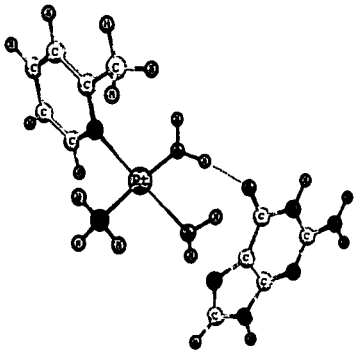


TS-3'

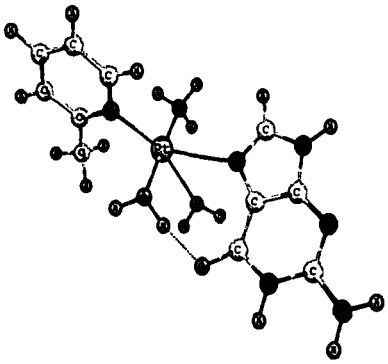


PC-3'

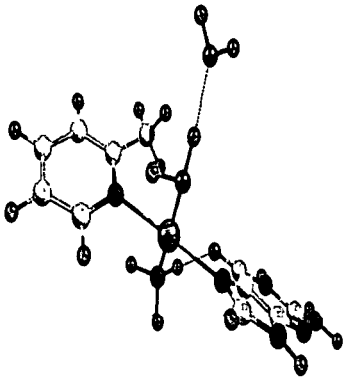
Step 4



RC-4



TS-4



PC-4

Figure 6.4: Stationary points (RC, TS, and PC) along the reaction coordinate of guanine substitution reactions optimized at HF/6-31G* level of theory.

Substitution of aqua ligand *trans* to NH_3 in step 3' proceeds *via* a less stable transition state, TS-3'. The reactant complex (RC-3') obtained in the displacement of aqua ligand *trans* to amine by N7 of guanine is hydrogen bonded through H_2O ligand and O6 atom of guanine molecule. Since hydrogen bonds formed between different groups are not of equal strength, the reactant complexes in these two steps are not equally stable. The stability of RC-3 is 5.6 kcal/mol higher than RC-3'. In PC-3 structure, water molecule is hydrogen bonded with chlorine atom and also with guanine hydrogen. Notable structural differences between the transition state and product complex in step 3' is observed where the water molecule shift to a different position making strong hydrogen bonds with O6 atom (2.11Å) and amine hydrogen atom (2.08Å), respectively.

Although we have observed that the most probable species interacting with DNA are mono-charged complexes of *cis*-$[PtCl_2(NH_3)(2\text{-picoline})]$, binding of guanine with diaquated form of the drug is also considered as increasing positive charges of metal species increases the affinity of metal complexes.[46] In this investigation (for step 3 and step 4), only possible hydrogen bonding between amine and O6 atom of guanine is considered as it is prerequisite to have at least one N–H group bound to O6 atom of guanine for the anticancer activity of drug.[47]

In step 4, the water molecule *trans* to picoline group of the diaquo species bearing two positive charges is replaced by guanine. In RC-4 structure, one hydrogen bond connects between N7 atom of guanine with aquo ligand and the other one connects O6 atom of guanine with amine ligand. The shape of the transition state, TS-4, is similar with the other transition states having hydrogen bond between amine group and O6 atom. The entering guanine ligand approaches the reaction centre with a longer bond length (2.65Å) than the leaving aquo ligand (2.48Å). The PC-4 has hydrogen bond between leaving water group and amine ligand. For all steps, corresponding structures of reactants and products are confirmed by IRC calculations.

6.3.2.2 Kinetic Analysis

The Gibbs free energies and rate constants calculated at HF/6-31G* level for binding of the hydrated drug molecule with guanine in gas phase as well as solvent phase are presented in Table 6.4.

Table 6.4: Gibbs free energies and rate constants for three steps of substitutions with guanine. $\Delta G^{\ddagger}$ values are in kcal mol^{-1}. Solvent phase values are in parenthesis.

	HF/631-G*	Expt.[a]
$\Delta G_3^{\ddagger}$	21.33 (19.15)	
$\Delta G_{-3}^{\ddagger}$	26.55	
$\Delta G_{3'}^{\ddagger}$	22.82 (20.54)	
$\Delta G_{-3'}^{\ddagger}$	25.86	
$\Delta G_4^{\ddagger}$	20.33 (19.04)	
$\Delta G_{-4}^{\ddagger}$	26.76	
k_3	1.43×10^{-3} (5.67×10^{-2})	6.67×10^{-3}
$k_{3'}$	1.11×10^{-4} (5.42×10^{-3})	7.97×10^{-3}
k_4	7.63×10^{-3} (6.83×10^{-2})	

[a] From Ref. 45

The inclusion of solvation effect by means of continuum PCM model reduces the activation energy from that of gas phase values. The step 3 needs lower barrier of about 21.3 kcal/mol in gas phase than the energy barrier needed for step 3' (22.8 kcal/mol). The solvent phase value for step 3 is also lower than step 3'. The replacement of aqua ligand by guanine from doubly charged reactant complex in step 4 needs the lowest barrier of about 20.3 kcal/mol (19.0 kcal/mol in water). This result is different from those obtained for cispatin by Chval *et al.*[19] They obtained highest activation barrier for binding of guanine with diaquo species compared to the mono aquo species. However, our results on hydrolysis reactions predict that mono aquo products are the probable species that interact with guanine. Thus B (Scheme 6.1) is the most probable species to attack guanine according to our results. Moreover, the reactant complex, RC-3 includes large portion of electrostatic

interaction that seems to be responsible for stronger stabilization of guanine adduct obtained *via* step 3 when compared with step 3'. Figure 6.5 shows the reaction pathways obtained for binding of guanine with mono aquo (Figure 6.5. a) and with diaquo species (Figure 6.5. b).

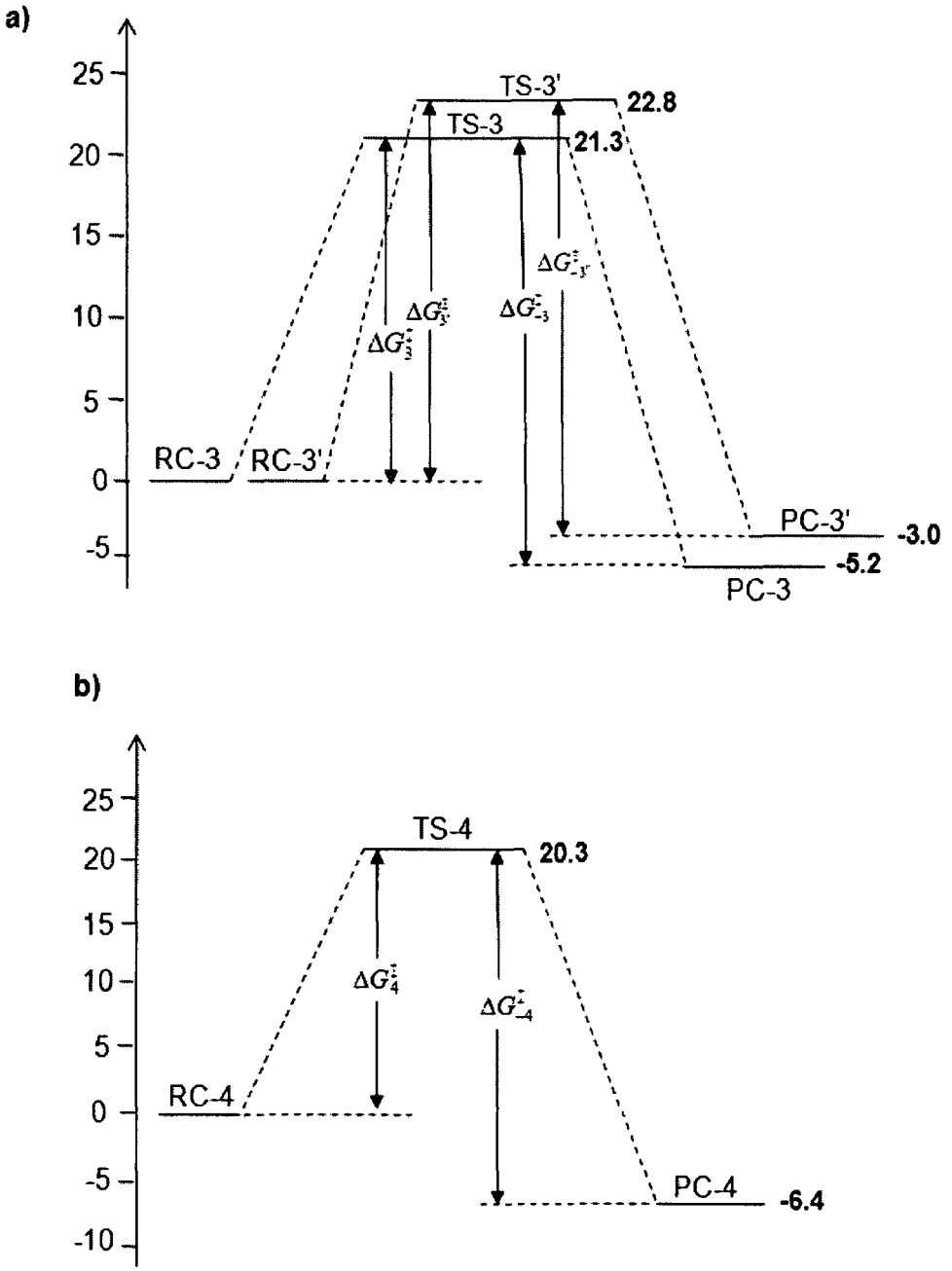


Figure 6.5. Potential energy profile for guanine binding reactions with monoaquo (a) and diaquo (b) products of AMD473 at HF/6-31G* level of theory.

All reactions are exothermic, step 3 = -5.2 kcal/mol, step 3' = -3.0 kcal/mol, and step 4 = -6.4 kcal/mol. The rate constants for all ligand exchange reactions with guanine for both gas phase and solvent phase are presented in Table 6.4. The values for k_3 (1.43×10^{-3} $s^{-1}M^{-1}$) calculated at gas phase are in good agreement with the experimental value (6.67×10^{-3} $s^{-1}M^{-1}$).[45] However, its value in solution differs by one order of magnitude compared to the experimental value. This may be due to lack of DNA backbone which would provide steric hindrance and decrease the solvent effect. The rate constant $k_{3'}$ in solution (5.42×10^{-3} $s^{-1}M^{-1}$), on the other hand agrees well with the available experimental value (7.97×10^{-3} $s^{-1}M^{-1}$).[45] The higher values of rate constants for k_4 in both gas and solvent media compared to the other two rate constants accounts for higher affinity of the doubly charged species towards guanine. Inconsistent with the experimental observation[45], we found higher values of k_3 over $k_{3'}$ in both gas and solvent phases (Table 6.4). The substitution of aqua ligand *cis* to picoline by N7 of guanine (step 3') experience a steric hindrance provided by the axial 2-picoline ligand. Thus the higher values of k_3 over $k_{3'}$ is due to the well known fact that axial steric hindrance slows down the rate of ligand substitution reactions in square-planar metal complexes.

6.4 Conclusions

In the present work we have presented a systematic analysis of hydrolysis mechanism of anticancer drug AMD473 (*cis*-[$PtCl_2(NH_3)$(2-picoline)]) using HF and different density functional levels of theory. The monofunctional binding of aquo species with guanine are also performed with HF theory to provide a detailed binding mechanism of the anticancer drug. The structures of all species on the reaction path are confirmed by using IRC calculations. The geometries of the stationary points of these ligand substitution reactions agree well with that of the parent compound cisplatin. Four different steps of hydrolysis reactions are considered and energy barriers as well as rate constants are calculated for each process at all levels of theory. The gas phase calculated rate constants for first hydrolysis reactions closely agree with the experimental data. For the second step the gas phase theoretical values are quite distinct from the experimental values which become comparable with the incorporation of solvent environment using

polarizable continuum model (PCM). The activation barrier for second step is higher than the first step and thus monoaquo forms are the most probable species that react with DNA. Slightly lower activation energy accompanied with higher stability of the stationary points in step 1 (hydrolysis of Cl^- *trans* to 2-picoline) over step 1' (hydrolysis of Cl^- *trans* to amine) makes the former path more favorable. The transition states predicted in ligand exchange reactions of mono and diaquo forms of AMD473 with guanine have pentacoordinated geometry as expected for associative mechanism of square planar structure. Formation of the reactant complexes is stabilized by mainly hydrogen bond connecting amine group and O6 atom of guanine which is supported by experimental observation. The activation energy for step 3' is slightly higher by about 2.0-3.0 kcal/mol compared to step 3 and step 4 in both gas and solvent phases which is due to the steric hindrance experienced by the axial 2-picoline group. The step 4 needs the lowest barrier (19.0 kcal/mol) in solution to proceed. However, being monoaquo forms as preferred species, step 3 is the most probable path for guanine binding. Products (D and D', Scheme 6.1) obtained from the two step reactions: initial hydrolysis followed by guanine substitution, will undergo further hydrolysis and guanine binding to give bifunctional DNA-Pt adduct. Thus as a result of steric effect and asymmetric structure of AMD473, different forms of guanine adduct of the drug are possible.

References

(1) Rosenberg, B.; Van Camp, L.; Krigas, T. Inhibition of cell division in Escherichia coli by electrolysis products from a platinum electrode. *Nature* **205**, 698–699 (1965)

(2) Rosenberg, B.; Van Camp, L.; Trosko, J.E.; Mansour, V.H. Platinum compounds: A new class of potent antitumour agents. *Nature* **222**, 385–386 (1969)

(3) Jamieson, E.R.; Lippard, S.J. Structure, recognition, and processing of cisplatin–DNA adducts. *Chem. Rev.* **99**, 2467–2498 (1999)

(4) Wong, E. Giandomenico, C.M. Current status of platinum-based antitumor drugs. *Chem. Rev.* **99**, 2451–2466 (1999)

(5) Reedijk, J. New clues for platinum antitumor chemistry: Kinetically controlled metal binding to DNA. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3611–3616 (2003)

(6) Wang, D.; Lippard, S.J. Cellular processing of platinum anticancer drugs. *Nat. Rev. Drug Discovery* **4**, 307–320 (2005)

(7) Holford, J. et al. Chemical, biochemical and pharmacological activity of the novel sterically hindered platinum co-ordination complex, cis-[amminedichloro(2-methylpyridine)] platinum(II) (AMD473). *Anti-Cancer Drug Des.* **13**, 1–18 (1998)

(8) Sharp, S.Y.; O'Neill, C.F.; Rogers, P.; Boxall, F.E.; Kelland, L.R. Retention of activity by the new generation platinum agent AMD0473 in four human tumour cell lines possessing acquired resistance to oxaliplatin. *Eur. J. Cancer.* **38**, 2309–2315 (2002)

(9) Chen, Y.; Guo, Z.; Parsons, S.; Sadler, P.J. Stereospecific and kinetic control over the hydrolysis of a sterically hindered platinum picoline anticancer complex. *Chem. Eur. J.* **4**, 672–676 (1998)

(10) Wysokinski, R.; Michalska, D. The Performance of different density functional methods in the calculation of molecular structures and vibrational spectra of platinum(II) antitumor drugs: Cisplatin and carboplatin. *J. Comput. Chem.* **9**, 901–912 (2001)

(11) Wysokinski, R.; Michalska, D. The prediction of Raman spectra of platinum(II) anticancer drugs by density functional theory. *Chem. Phys. Lett.* **403**, 211–217 (2005)

(12) Chval, Z.; Sip, M. Pentacoordinated transition states of cisplatin hydrolysis–ab initio study. *J. Mol. Struct. THEOCHEM* **532**, 59–68 (2000)

(13) Zhang, Y.; Guo, Z.; You, X.-Z. Hydrolysis theory for cisplatin and its analogues based on density functional studies. *J. Am. Chem. Soc.*, **123**, 9378–9387 (2001)

(14) Costa, L.A.S.; Rocha, W.R.; DeAlmeida, W.B.; Santos, H.F.D. The hydrolysis process of the cis-dichloro.ethylenediamine.platinum.II.: A theoretical study. *J. Chem. Phys.* **188**, 10584–10592 (2003)

(15) Raber, J. Zhu, C.; Eriksson, L.A. Activation of anti-cancer drug cisplatin — is the activated complex fully aquated? *Mol. Phys.* **102**, 2537–2544 (2004)

(16) Burda, J.V.; Zeizinger, M.; Leszcsynski, J. Activation barriers and rate constants for hydration of platinum and palladium square-planar complexes: An ab initio study. *J. Chem. Phys.* **120**, 1253–1262 (2004)

(17) Robertazzi, A. Platts, J.A. Hydrogen bonding, solvation, and hydrolysis of cisplatin: A theoretical study. *J. Comput. Chem.* **25**, 1060–1067 (2004)

(18) Lau, J. K-C.; Deubel, D.V. Hydrolysis of the anticancer drug cisplatin: Pitfalls in the interpretation of quantum chemical calculations. *J. Chem. Theory. Comput.* **2**, 103–106 (2006)

(19) Chval, Z.; Sip, M. Transition states of cisplatin binding to guanine and adenine: *ab initio* reactivity study. *Collect. Czech. Chem. Commun.* **68**, 1105–1118 (2003)

(20) Raber, J.; Zhu, C.; Eriksson, L.A. Theoretical study of cisplatin binding to DNA: The importance of initial complex stabilization. *J. Phys. Chem. B* **109**, 11006–11015 (2005)

(21) Costa, L.A.S.; Hambley, T.W.; Rocha, W.R.; DeAlmeida, W.B.; Santos, H.F.D. Kinetics and structural aspects of the cisplatin interactions with guanine: A quantum mechanical description. *Int. J. Quantum. Chem.* **106**, 2129–2144 (2006)

(22) Burda, J.V.; Leszczynski, J. How strong can the bend be on a DNA helix from cisplatin? DFT and MP2 quantum chemical calculations of cisplatin-bridged DNA purine basese. *Inorg. Chem.* **42**, 7162–7172 (2003)

(23) Zeizinger, M.; Burda, J.V.; Leszczynski, J. The influence of a sugar-phosphate backbone on the cisplatin-bridged BpB' models of DNA purine bases.

Quantum chemical calculations of Pt(II) bonding characteristics. *J. Phys. Chem. Chem. Phys.* **6**, 3585–3590 (2004)

(24) Matsui, T.; Shigeta, Y.; Hirao, K. Influence of Pt comlex binding on the guanine-cytosine pair: A theoretical study. *Chem. Phys. Lett.* **423**, 331–334 (2006)

(25) Matsui, T.; Shigeta, Y.; Hirao, K. Multiple proton-transfer reactions in DNA base pairs by coordination of Pt complex. *J. Phys. Chem. B* **111**, 1176–1181 (2007)

(26) Carloni, P.; Sprik, M.; Adreoni, W. Key steps of the cisplatin-DNA interaction: Density functional theory-based molecular dynamics simulations. *J. Phys. Chem. B* **104**, 823–835 (2000)

(27) Deubel, D.V. Factors governing the kinetic competition of nitrogen and sulfur ligands in cisplatin binding to biological targets. *J. Am. Chem. Soc.* **126**, 5999–6004 (2004)

(28) Deubel, D.V. The chemistry of dinuclear analogues of the anticancer drug cisplatin. A DFT/CDM study. *J. Am. Chem. Soc.* **128**. 1654–1663 (2006)

(29) Frisch, M.J. *et al Gaussian 03;* Revision D01 Gaussian, Inc.: Wallingford, CT, 2004

(30) Pavankumar, P.N.V. *et al* Comprehensive *ab initio* quantum mechanical and molecular orbital (MO) analysis of cisplatin: Structure, bonding, charge density, and vibrational frequencies. *J. Comput. Chem.* **20**, 365–382 (1999)

(31) Perdew, J.P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **33**, 8822–8824 (1986)

(32) Perdew, J.P. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* **46**, 6671–6687 (1992)

(33) Perdew, J.P.; Burke, K.; Wang, Y. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys. Rev. B* **54**, 16533–16539 (1996)

(34) Adamo, C.; Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The *m*PW and *m*PW1PW models. *J. Chem. Phys.* **108**, 664–675 (1998)

(35) Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996)

(36) Hay, P.J.; Wadt, W.R. *Ab initio* effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.* **82**, 299 (1985)

(37) Ditchfield, R.; Hehre, W.J. Pople, J. A. Self-consistent molecular-orbital methods. IX. An extended gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **54**, 724–728 (1971)

(38) Gonzalez, C.; Schlegel, H.B. An improved algorithm for reaction path following. *J. Chem. Phys.* **90**, 2154–2162 (1989)

(39) Gonzalez, C.; Schlegel, H.B. Reaction path following in mass-weighted internal coordinates. *J. Chem. Phys.* **94**, 5523–5527 (1990)

(40) Cancès, M.T.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **107**, 3032–3042 (1997)

(41) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. Ab initio study of ionic solutions by a polarizable continuum dielectric model. *Chem. Phys. Lett.*, **286**, 253–260 (1998)

(42) Mennucci, B.; Tomasi, J. Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys.* **106**, 5151–5159 (1997)

(43) Costa, L.A.S.; Rocha, W.R.; DeAlmeida, W.B.; Santos, H.F.D. The solvent effect on the aquation processes of the *cis*-dichloro(ethylenediammine)platinum(II) using continuum solvation models. *Chem. Phys. Lett.* **387**, 182–187 (2004)

(44) Michalska, D.; Wysokinski, R. Molecular structure and bonding in platinum-picoline anticancer complex: Density functional study. *Collect. Czech. Chem. Commun.* 69 63–72 (2004)

(45) Chen, Y.; Guo, Z.; Parikinson, J. A.; Sadler, P.J. Kinetic control of reactions of a sterically hindered platinum picoline anticancer complex with guanosine 5'-monophosphate and glutathione *J. Chem. Soc., Dalton Trans.* 3577–3585 (1998)

(46) Black, C.B.; Cowan, J.A. Quantitative evaluation of electrostatic and hydrogen bonding contributions to metal cofactor binding to nucleic Acids. *J. Am. Chem. Soc.* **116**, 1174–1178 (1994)

(47) Reedijk, J. Improved understanding in platinum antitumour chemistry. *Chem. Commun.* 801–806 (1996)

Chapter 7

Stability and Proton Transfer in DNA Base Pairs of AMD473-DNA Adduct

In this chapter, we investigate the energetics as well as structural differences of the four different adducts of cisplatin analogue *cis*-[$PtCl_2(NH_3)$(2-picoline)] (AMD473) with a duplex DNA calculated by using DFT and quantum mechanics/molecular mechanics (QM/MM) based two layer ONIOM methods to probe their stabilities. Further, we have studied the possibilities of proton transfer between DNA base pairs of the most stable drug-DNA adduct which leads to mutation of the DNA. The adduct **b** (2-picoline *trans* to 3'G and 2-methyl group directed to the DNA major groove center) is found to be the most stable configuration among all the possible adducts. From the proton transfer analysis we found that the single proton transfer between N1 position of guanine (G) and N3 position of cytosine (C) of each GC pair gives a structure energetically as stable as the original one. No simultaneous single proton transfer in both the pairs is energetically found to be stable. [Sarmah and Deka, *communicated*]

7.1 Introduction

Cisplatin,[1] is one of the most widely used anticancer drugs and is particularly active in treating several kinds of cancer, such as testicular and ovarian cancers. Although cisplatin is one of the most successful anticancer drugs, its toxic side effects, intrinsic and acquired cellular resistance and limited solubility in aqueous solution have motivated searches for structurally and functionally analogues alternatives. In this way, second- and later third-generation drugs (like carboplatin, oxaliplatin, and dinuclear or trinuclear species) were discovered. Unfortunately, these drugs also suffer from resistance and other side effects.[2,3] One new promising anticancer agent, *cis*-[$PtCl_2(NH_3)$(2-picoline)], known as AMD473, has now entered the worldwide phase II and III clinical trials.[4] This orally administrated drug is less toxic than cisplatin and possesses activity against cisplatin-resistant cell lines. Specifically, the N7 atom of purine bases is the main binding site, with guanine being preferred over adenine. Indeed, as water is more labile than chloride, the reactive form of these molecules are believed to be the aqua species, which results from substitutions of one or both the chloride leaving groups by water molecules. Among several possible adducts, GpG adducts being the major and ApG cross-links being the next most abundant products.

Over the last few years, considerable theoretical efforts have been focused on cisplatin-DNA interaction to provide detailed insight into the binding mechanism at molecular level. There is a large amount of modelling studies on electronic structure and spectral analyses[5,6], structure-activity studies[7-9], aquation processes[10-14] structural properties of cisplatin-DNA complexes[15-19], effect on DNA base pairing[20,21], and chemical reactions responsible in developing toxic side-effects and resistance[22,23] of cisplatin and congeners. Despite all the effort in understanding cisplatin reactions and to some extent carboplatin[24] and oxaliplatin[25], a very few studies are devoted on binding mechanism of sterically hindered drug AMD473[26,27] and its interaction with DNA[28].

Due to the asymmetric structure of AMD473, it can form four stereoisomers with DNA. In this study we have investigated the stabilities of these AMD473-DNA adducts by discussing the energies and structural differences between these adducts in detail. Further we have explored the possibilities of proton transfer between DNA bases of the most stable drug-DNA adduct.

7.2 Methodology

The calculations are based on the experimental NMR structure[29] (PDB: 1A84) 5′d ($C_1C_2T_3C_4T_5G_6$*G_7*$T_8C_9T_{10}C_{11}C_{12}$) 3′d ($G_{24}G_{23}A_{22}G_{21}A_{20}C_{19}C_{18}A_{17}G_{16}$-$A_{15}G_{14}G_{13}$) in which we have manually replaced amine ligand with 2-picoline group. The NMR solution structure of AMD473 with DNA is the A.T rich 14-mer DNA duplex.[30] We have used the cisplatin-DNA duplex calculated from the NMR data[29], as it is reported that this duplex has essentially the same structural characteristics as the platinated 14-mer DNA structure.[30] Four different models of drug-DNA adducts are, **a** (2-picoline *trans* to 3′G_7 and 2-methyl group directed to the DNA backbone), **b** (2-picoline *trans* to 3′G_7 and 2-methyl group directed to the DNA major groove center), **c** (2-picoline *trans* to 5′G_6 and 2-methyl group directed to the DNA backbone) and **d** (2-picoline *trans* to 5′G_6 and 2-methyl group directed to the DNA major groove center). All the structures were solvated by 600 water molecules. Sodium counter ions were used to balance the DNA backbone charges. We adopted QM/MM based ONIOM method to optimize the final configurations of 4 adducts where QM includes *cis*-$[Pt(NH_3)(2\text{-picoline})]^{2+}$ part, G_6 and G_7 while the remaining DNA bases, sugar-phosphate backbone and water molecules were treated with UFF. All the structures were optimized by HF and DFT methods with BHandH[31] and mPW1PW91[32] functionals using Gaussian 03 program.[33] We have used LanL2DZ basis set as described by Hay and Wadt[34] to treat Pt whereas all other atoms of the complexes were treated at 6-31G(d,p) level. In proton transfer calculations QM part has been extended upto two base pairs (G_6C and G_7C) and the other atoms were treated with lower level.

7.3 Results and Discussion

The energies obtained by QM/MM method for all configurations calculated at different levels of theory are presented in Table 7.1. According to the results obtained with all studied theoretical methods, we found the model **b** as lowest energy structure. The *cis*-$[PtCl_2(NH_3)(2\text{-picoline})G^*G^*]$ (QM part) moiety of the adduct **b** has the lowest energy as shown in the Table7.1 which suggests the adduct to have the most stable configuration. In agreement with our results, the thermodynamic and kinetic studies of Chen *et al.*[30] also favor the formation **b**. This is mainly due to the steric selection of picoline ring which fits perfectly into the

major groove of the DNA duplex. It is observed that there is an energy difference of 9.4 kcal/mol between **a** and **b** at BHandH level whereas the model **b** is almost 52.6 kcal/mol and 22.5 kcal/mol stable than **c** and **d**, respectively obtained at the same level. Despite the energy differences, different level of theories predict similar trend of stability as **b** > **a** > **d** > **c.**

Table 7.1: The QM/MM optimized energies (a.u) for the four adducts. QM energies are shown in parenthesis.

	QM/MM Energy			
	a	**b**	**c**	**d**
HF/[LANL2DZ-ECP + 6-31G**]	-1538.352	-1538.583	-1538.217	-1538.280
	(-1537.808)	(-1537.812)	(-1537.801)	(-1537.805)
BHandH/[LANL2DZ-ECP + 6-31G**]	-1539.480	-1539.495	-1539.411	-1539.459
	(-1538.660)	(-1538.663)	(-1538.651)	(-1538.658)
mPW1PW91/[LANL2DZ-ECP + 6-31G**]	-1546.651	-1546.722	-1546.606	-1546.622
	(-1547.558)	(-1547.563)	(-1547.554)	(-1547.555)

Although the energies calculated at mPW1PW91/[LANL2DZ-ECP + 6-31G**] level have the lowest values, all the results along with the proton transfer analysis are discussed at BHandH/[LANL2DZ-ECP + 6-31G**] level, as BHandH functional is shown to be efficient for the study of л-stacked systems like DNA. The binding energies for the bifunctionals adduct of AMD473 with DNA, along with other geometrical parameters are presented in the Table 7.2. From the calculated binding energies, we found **b** as the most stable structure with the highest value. It is seen that the model **a** with binding energy 7 kcal/mol lower, being the next stable

structure. Comparing the values, the binding energy of the less stable structures **c** and **d** are about 31 kcal/mol and 25 kcal/mol lower than that of the structure **b**, respectively.

Table 7.2: Binding energies (BE) and optimized geometries* of the four adducts.

	a	**b**	**c**	**d**
BE (kcal/mol)	742.47	749.36	717.82	723.96
$Pt–N7_{amine}$	2.03	2.04	2.03	2.03
$Pt–N7_{pico}$	2.02	2.01	2.01	2.02
$Pt–N7G_6$	2.04	2.01	2.02	2.02
$Pt–N7G_7$	2.03	2.02	2.03	2.04
$N7G_6–Pt–N7G_7$	85.3	89.0	86.4	85.4
$N7_{amine}–Pt–N7_{pico}$	88.6	91.5	91.3	88.0

* The unit of distance is angstroms, the unit of angle is degrees.

Figure 7.1 shows the optimized structures of four possible adducts of AMD473 drug with DNA. The picoline ring forms $(H_3C)C–N$(picoline)–Pt–N(ammonia) dihedral angle equal to 111.7° in model **a**. The deviation of the picoline ligand from the perpendicular orientation with the Pt square plane shortens the distance between amine and the 2-methyl of the picoline group. The angles between Pt square plane and the planes of G_6 and G_7 are 61° and 73.2°, respectively.

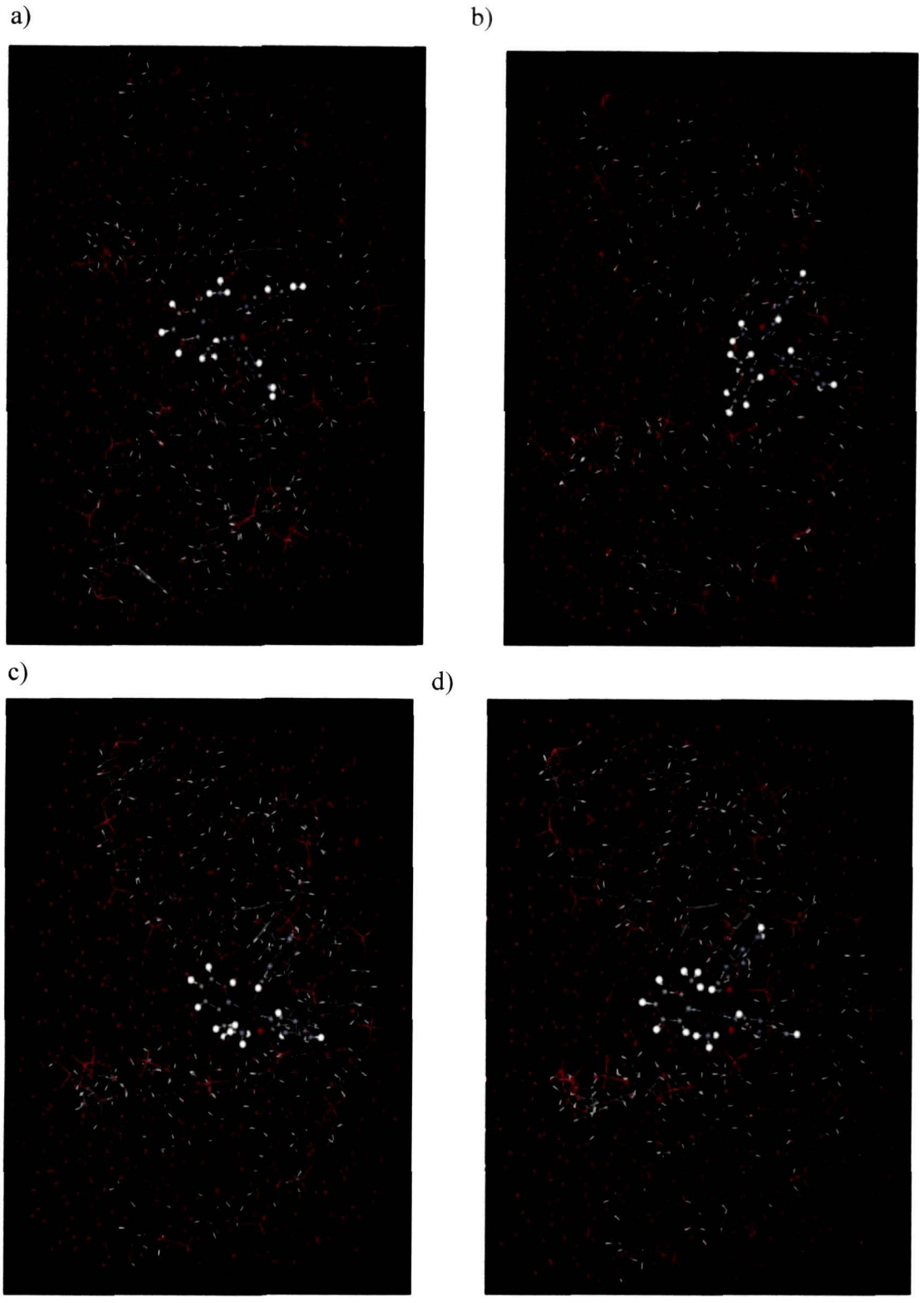


Figure 7.1: Structures of the AMD473 drug-DNA adducts optimized by QM/MM method. Drug, G_6, and G_7 belong to the QM region (in balls and sticks), the rest of the DNA and the solvent (only oxygen atoms are shown for clarity) (in lines) belong to the MM region.

In **b,** the picoline ring fits perfectly into the DNA major groove and lies along the phosphate backbone. This structural arrangement is in accordance with the experimental observation[30], as shown in the Figure 7.2.

i) ii)

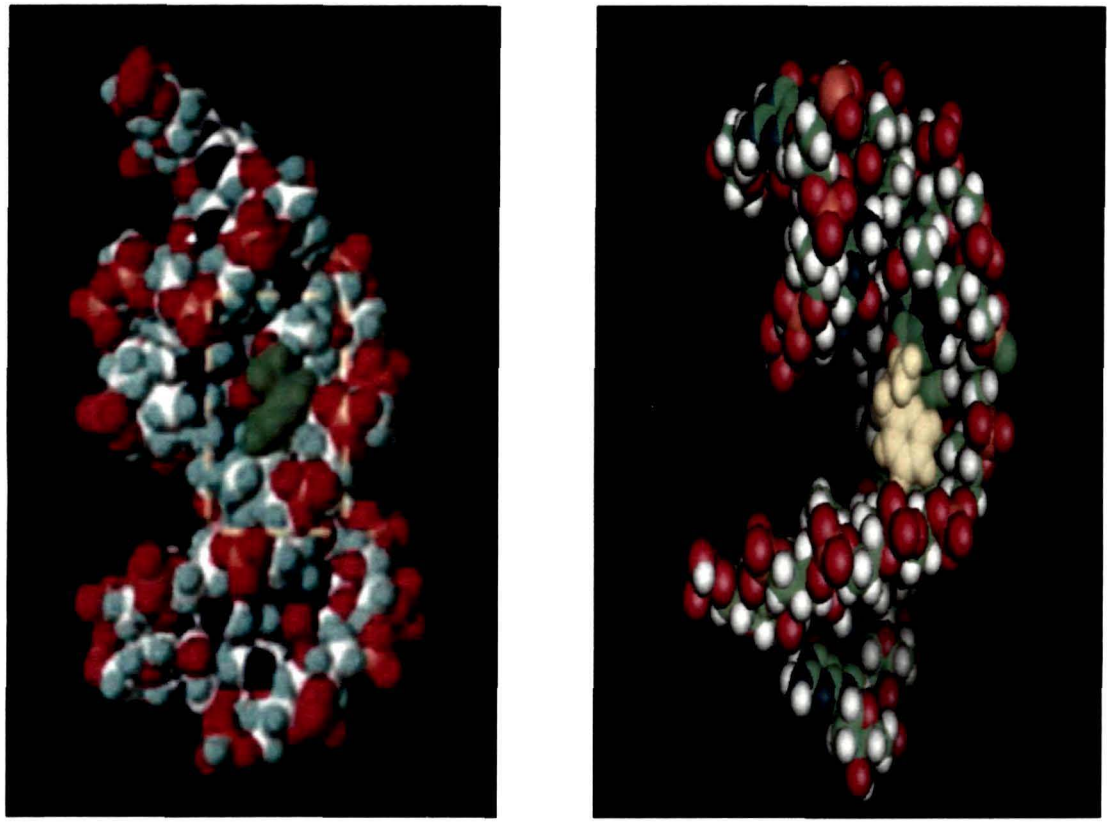


Figure 7.2: Space filling model of experimental i) and optimized ii) geometries of AMD473–DNA adduct (**b**) with *cis*-$[Pt(NH_3)(2\text{-Picoline})]^{2+}$ is shown in green and yellow, respectively.

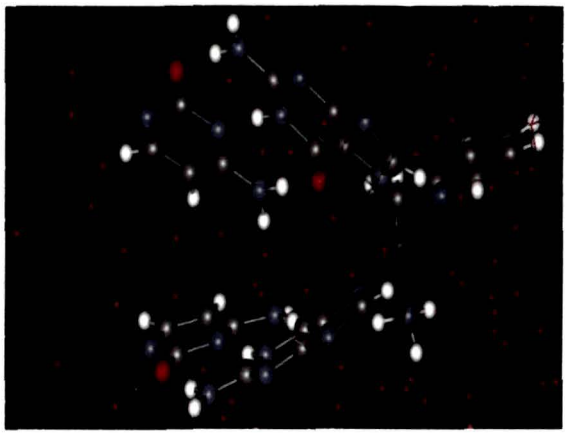
The picoline ring is found to be perpendicular to the molecular plane with the dihedral angle equal to 102.3° and H_3C– Pt distance equal to 3.19 Å in agreement with previous theoretical and experimental results.[35,36] The angle between picoline ring and G_7 is 57.2° whereas it forms an angle of 77.8° with G_6. The distance between a hydrogen atom of the amine and the $O6G_7$ is 1.91 Å and N-H-O is 150.3°, which indicates that a hydrogen bond exist, indeed. This gives an additional stability to the model. The Pt-N7G bond is also slightly shorter and stronger in **b** compared to that of the other structures (Table 7.2). The steric hindrance provided by the axial

2-picoline ligand in **b** is more prominent than that in others and thus the activity against cisplatin-resistant cell lines is more effective. The high stability of **b** may be a reason behind the lack of cross-resistance between AMD473 and cisplatin. In **c,** the angle between the Pt square plane and picoline ligand is larger (113.5°) than that of **a** and **b** and this increases the distance between amine and the 2-methyl of picoline group. Model **d** has the angle of 117.6° between the Pt square plane and picoline ring. The 2-picoline ligand forms an angle of ~90° with the planes of G_6 and G_7 bases in both models **c** and **d**. The perpendicular orientation of picoline ring to the planes of nearby bases, resulting in an unfavourable steric interactions which destabilizes the adducts. Thus from the observations, concerning both energy and structural properties of the four models, we found **b** as the most stable configuration.

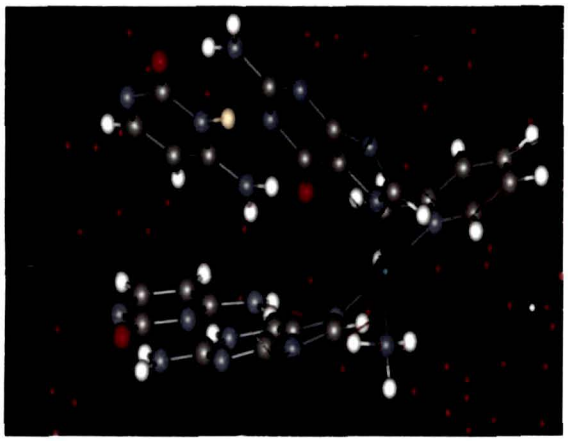
Further, we have investigated the energetic of intermolecular proton transfer (PT) in Watson-Crick base pairs by considering the most stable drug-DNA adduct (**b**) in water environment. This investigation may allow us to reveal a realistic picture of mispairing of base pairs which leads to the mutation of DNA. We observed that the adduct undergoes single proton transfer between N1 (G) and N3 (C) while the simultaneous single proton transfer in two stacked base pairs is not found to be stable. This is in agreement with the previous study on 4 base pairs model of cisplatin guanine adduct.[21] Figure 7.3 shows the optimized structures of **(1)** *cis* (G_6C)–Pt–(G_7C), **(2)** *cis* ($G_6^{PT}C$)–Pt–(G_7C), and **(3)** *cis* (G_6C)–Pt–($G_7^{PT}C$), where G_6^{PT} and G_7^{PT} denotes the proton donor guanine.

The drug binding changes the geometries of all the bases in QM region. This is expected from previous studies on smaller model complexes, which report strong perturbation of the base-pairing by cation binding. Although the G_7C pair slightly deviates from planarity, G_6C pair is greatly distorted. This may be due to the *cis* orientation of G_6C pair with respect to the bulky 2-picoline ring. Table 7.3 summarizes the hydrogen bond distances of all the optimized structures and compares them with the experimental data of base pairs (NMR structure, PDB: 1A84). The distance between O_6 (G) and N_4 (C) reduces by about 0.09 Å and 0.15 Å in the proton transferred GC pairs of (2) and (3), respectively, in comparison with the original GC pairs (1). The hydrogen bonding length between N_1 (G) and N_3 (C) also reduces whereas this length between N_2 (G) and O_2 (C) increases. Similar trends have been observed in the previous theoretical study.[37]

1)



2)



3)

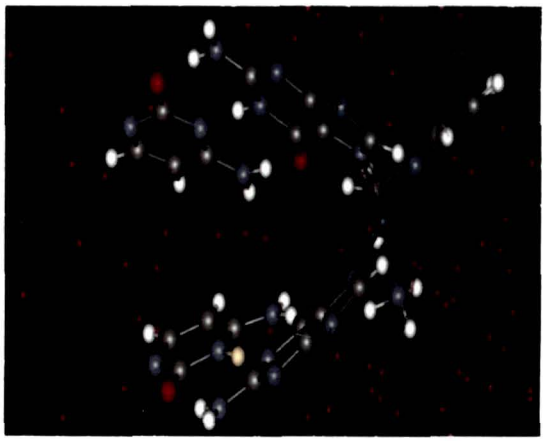


Figure 7.3: The optimized structures of (1) *cis*-(G_6C)–Pt–(G_7C), (2) *cis*-($G_6^{PT}C$)–Pt–(G_7C), and (3) *cis*-(G_6C)–Pt–($G_7^{PT}C$) of AMD473-DNA adduct. The transferred proton is represented by yellow colour.

Table 7.3: Hydrogen bond distances (Å) of experimental and computed non proton transferred and single proton transferred structures.

	cis-(G_6C)–Pt–(G_7C)		*cis*-($G_6^{PT}C$)–Pt–(G_7C)		*cis*-(G_6C)–Pt–($G_7^{PT}C$)		1A84[29]	
	G_6C	G_7C	G_6C	G_7C	G_6C	G_7C	G_6C	G_7C
O_6–N_4	2.79	2.84	2.70	2.87	2.86	2.69	2.90	2.90
N_1–N_3	2.8	2.82	2.70	2.83	2.83	2.74	2.53	2.93
N_2–O_2	2.72	2.69	2.81	2.70	2.71	2.89	2.34	2.84

To analyze the changes produced on the charge distribution of the QM part upon proton transfer, we present the sum of the NBO (Natural Bond Orbital) charges of the bases and the ligands in Table 7.4. The QM regions are extracted from their corresponding overall structures, link atoms replaced with hydrogen and then single point calculations are carried out to calculate the NBO charges.

Table 7.4: The sum of NBO charges on bases and ligands.

	cis (G_6C)–Pt–(G_7C)	*cis* ($G_6^{PT}C$)–Pt–(G_7C)	*cis*(G_6C)–Pt–($G_7^{PT}C$)
G_6 or G_6^{PT}	0.211	-0.491[a]	0.213
G_7 or G_7^{PT}	0.327	0.231	-0.466[a]
$C(G_6)$	0.054	0.782[b]	0.080
$C(G_7)$	0.008	0.103	0.807[b]
NH_3	0.273	0.24411	0.248
2-picoline	0.280	0.26127	0.253

[a] The sum does not contain transferred H atom.

[b] The sum contains transferred H atom.

The guanine bases in (1) are positive whereas the cytosine bases remained almost neutral. The positive value of G is mainly due to the charge transfer from the bases to Pt atom of the drug. After proton transfer, the sum of the charges on G in (2) and (3) becomes negative. On the other hand, the entire charge on C in PT products has a positive value which causes a charge separation between G and C leading to a stable $G^{PT}C$ pair due to Coulomb attraction between them. The simultaneous single proton transfer has been prevented since the two stacked guanine bases become negatively charged and both the cytosine bases bear positive charge which results Coulomb repulsions among the bases.

The energetic details of the proton transferred products are given in Table 7.5. We observed that the single proton transfer between G and C in which the guanine molecule is *trans* to 2-picoline group (3) is energetically more favorable

than that of *trans* to amine (2). This difference is may be due to their difference in planarity. The *cis* (G_6C)–Pt–(G_7C) and *cis*(G_6C)–Pt–($G_7^{PT}C$) pairs are almost equal in energy. The energy differences between the products (2 and 3) and the original structure (1) are about 2–1.3 kcal/mol.

Table 7.5: Energetic of the PT products (kcal/mol).

cis (G_6C)–Pt–(G_7C)	0
cis ($G_6^{PT}C$)–Pt–(G_7C)	+2.3
cis(G_6C)–Pt–($G_7^{PT}C$)	+1.5

7.4 Conclusions

Our calculations analyze the stabilities of four stereoisomers formed by binding of AMD473 with DNA. The orientations of picoline group with respect to the planes of two platinated guanine bases destabilizes the adducts **c** and **d** compared to **a** and **b**. The picoline plane is almost perpendicular to the Pt square plane in **b** and the picoline ligand perfectly fits into the DNA major groove. Also there is a favorable hydrogen bonding between NH_3 and O_6 of G_7 in model **b**. All these features enhance the stability of **b** than the other configurations. Also from the calculated energetics we found **b** as the most stable adduct. The conclusions obtained from intermolecular proton transfer reactions in DNA base pairs are that no simultaneous single proton transfer is energetically found to be stable while *cis*(G_6C)–Pt–($G_7^{PT}C$) pair is as stable as the original structure. The small difference in energy between the proton transferred products and the original one explains the influence of drug binding to induce DNA damage through base pair modification.

References

(1) Rosenberg, B.; Camp, L. V. Van,; Krigas, T. Inhibition of cell division in Escherichia coli by electrolysis products from a platinum electrode. *Nature*, **205**, 698–699 (1965)

(2) Reedijk, J. New clues for platinum antitumor chemistry: Kinetically controlled metal binding to DNA. *Proc. Natl. Acad. Sci. USA* **100**, 3611–3616 (2003)

(3) Wang, D.; Lippard, S. J. Cellular processing of platinum anticancer drugs. *Nat. Rev. Drug Discovery*, **4**, 307–320 (2005)

(4) Sharp, S. Y.; Neill, C. F. O.; Rogers, P.; Boxall, F. E.; Kelland,; L.R. Retention of activity by the new generation platinum agent AMD0473 in four human tumour cell lines possessing acquired resistance to oxaliplatin. *Eur. J. Cancer.* **38**, 2309–2315 (2002)

(5) Wysokinski, R.; Michalska, D.; The performance of different density functional methods in the calculation of molecular structures and vibrational spectra of platinum(II) antitumor drugs: Cisplatin and carboplatin *J. Comput. Chem.* **9**, 901–912 (2001)

(6) Michalska, D.; Wysokinski, R.; The prediction of Raman spectra of platinum(II) anticancer drugs by density functional theory. *Chem. Phys. Lett.* **403**, 211–217 (2005)

(7) Sarmah, P. ; Deka, R. C. Solvent effect on the reactivity of *cis*-platinum (II) complexes: A density functional approach. *Int. J. Quantum. Chem.* **108**, 1400–1409 (2008)

(8) Sarmah, P.; Deka, R. C. DFT-based QSAR and QSPR models of several cis-platinum complexes: solvent effect. *J. Comput. Aided. Mol. Des.* **23**, 343–354 (2009)

(9) Dans, P. D.; Coitiño, E. L. Density functional theory characterization and descriptive analysis of cisplatin and related compounds. *J. Chem. Inf. Model.* **49**, 1407–1419 (2009)

(10) Chval, Z.; Sip, M. Pentacoordinated transition states of cisplatin hydrolysis – ab initio study. *J. Mol. Struct. THEOCHEM*, **532**, 59–68 (2000)

(11) Zhang, Y.; Guo, Z.; You, X.-Z. Hydrolysis theory for cisplatin and its analogues based on density functional studies. *J. Am. Chem. Soc*, **123**, 9378–9387 (2001)

(12) Costa, L.A.S.; Rocha, W. R.; Almeida, W.B.D; Dos Santos, H.F. The hydrolysis process of the cis-dichloro.ethylenediamine.platinum.II.: A theoretical study. *J. Chem. Phys.* **188**, 10584–10592 (2003)

(13) Raber, J.; Zhu, C.; Eriksson, L.A. Activation of anti-cancer drug cisplatin — Is the activated complex fully aquated? *Mol. Phys.* **102**, 2537–2544 (2004)

(14) Lau, J.K-C.; Deubel, D.V. Hydrolysis of the anticancer drug cisplatin: Pitfalls in the interpretation of quantum chemical calculations. *J. Chem. Theory. Comput.* **2**, 103–106.(2006)

(15) Mantri, Y.; Lippard, S.J.; Baik, M.H. Bifunctional binding of cisplatin to DNA: Why does cisplatin form 1,2-intrastrand cross-links with AG but not with GA? *J. Am. Chem. Soc.* **129**, 5023–5030 (2007)

(16) Raber, J.; Zhu, C.; Eriksson, L.A. Theoretical study of cisplatin binding to DNA: The importance of initial complex stabilization. *J. Phys. Chem. B* **109**, 11006–11015 (2005)

(17) Zeizinger, M.; Burda, J.V.; Leszczynski, J. The influence of a sugar-phosphate backbone on the cisplatin-bridged BpB′ models of DNA purine bases. Quantum chemical calculations of Pt(II) bonding characteristics. *Phys. Chem. Chem. Phys.* **6**, 3585–3590 (2004)

(18) Robertazzi, A.; Platts, J.A. A QM/MM study of cisplatin–DNA oligonucleotides: From simple models to realistic systems. *Chem. Eur. J.* **12**, 5747–5756 (2006)

(19) Spiegel, K.; Rothlisberger, U.; Carloni, P. Cisplatin binding to DNA oligomers from hybrid Car-Parrinello/Molecular Dynamics simulations. *J. Phys. Chem. B* **108**, 2699–2707 (2004)

(20) Matsui, T.; Shigeta, Y.; Hirao, K. Influence of Pt complex binding on the guanine–cytosine pair: A theoretical study. *Chem. Phys. Lett.* **423**, 331–334 (2006)

(21) Matsui, T.; Shigeta, Y.; Hirao, K. Multiple proton-transfer reactions in DNA base pairs by coordination of Pt complex. *J. Phys. Chem. B* **111**, 1176–1181 (2007)

(22) Deubel, D.V. Factors governing the kinetic competition of nitrogen and sulfur ligands in cisplatin binding to biological targets. *J. Am. Chem. Soc.* **126**, 5999–6004 (2004)

(23) Deubel, D.V. The chemistry of dinuclear analogues of the anticancer drug cisplatin. A DFT/CDM study. *J. Am. Chem. Soc.* **128** 1654–1663 (2006)

(24) Pavelka, M.; Lucas, M.F.A.; Russo, N. On the hydrolysis mechanism of the second-generation anticancer drug carboplatin. *Chem. Eur. J.* **13**, 10108–10116 (2007)

(25) Lucas, M.F.A.; Pavelka, M.; Alberto, M.E.; Russo, N. Neutral and acidic hydrolysis reactions of the third generation anticancer drug oxaliplatin. *J. Phys. Chem. B*, **113**, 831–838 (2009)

(26) Sarmah, P.; Deka, R.C. Hydrolysis and binding mechanism of AMD473 (cis-[PtCl2(NH3)(2-picoline)]) with guanine: a quantum mechanical study. *J. Mol. Struct. THEOCHEM* **955**, 53–60 (2010)

(27) Banerjee, S.; Sengupta, P.S.; Mukherjee, A.K. A detailed theoretical DFT study of the hydrolysis mechanism of orally active anticancer drug ZD0473. *Chem. Phys. Lett.* **487**, 108–115 (2010)

(28) Jia, M.; Qu, W.; Yang, Z.; Chen, G. Theoretical study on the factors that affect the structure and stability of the adduct of a new platinum anticancer drug with a duplex DNA. *Int. J. Mod. Phys. B*, **19**, 2939–2949 (2005)

(29) Gelasco, A.; Lippard, S.J. NMR solution structure of a DNA dodecamer duplex containing a *cis*-diammineplatinum(II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry* **37**, 9230–9239 (1998)

(30) Chen, Y.; Parkinson, J.A.; Guo, Z.; Brown, T.; Sadler, P.J. A new platinum anticancer drug forms a highly stereoselective adduct with duplex DNA. *Angew. Chem. Int. Ed* **38**, 2060–2063 (1999)

(31) Becke, A.D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372–1378 (1993)

(32) Adamo, C.; Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The *m*PW and *m*PW1PW models. *J. Chem. Phys.* **108**, 664–675 (1998)

(33) Frisch, M.J.; *et al Gaussian 03;* Revision D01 Gaussian, Inc.: Wallingford, CT, 2004

(34) Hay, P.J.; Wadt, W.R. *Ab initio* effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.* **82**, 299–311 (1985)

(35) Michalska, D.; Wysokinski, R. Molecular structure and bonding in platinum-picoline anticancer complex: Density functional study. *Collect. Czech. Chem. Commun.* **69**, 63–72 (2004)

(36) Chen, Y.; Guo, Z.; Parsons, S.; Sadler, P.J. Stereospecific and kinetic control over the hydrolysis of a sterically hindered platinum picoline anticancer complex. *Chem. Eur. J.* **4**, 672–676 (1998)

(37) Noguera, M.; Bertran, J.; Sodupe, M. A quantum chemical study of Cu^{2+} interacting with guanine-cytosine base pair. Electrostatic and oxidative effects on intermolecular proton-transfer processes. *J. Phys. Chem. A* **108**, 333–341 (2004)

Chapter 8

Summary and Future Scopes

Cancer has a high mortality rate worldwide till date. It can occur as a result of mutation which is a consequence of DNA damage. There are different electronic properties of DNA which are responsible for its damage. Several approaches *viz.*, surgery, radiation therapy, chemotherapy, monoclonal antibody therapy or combinations of these therapies have been employed to treat cancer. The choice of appropriate treatment depends on the nature of tumor, the stage of the disease, and the general state of the patient. Since new anticancer agents have been discovered, the curing rate of chemotherapy has improved over the last decades. Organic drugs and natural products are most commonly used as chemotherapeutic agents. However, medicinal success of cisplatin, carboplatin, oxaliplatin and others unfold the possibility of transition metal compounds in cancer treatment. The toxicity and resistance factors are of significant importance, because conventional platinum drugs exert severe toxicity. Designing and development of new drugs from the laboratory bench to the clinic is a complex, expensive, and very time consuming task. In this context rational computer-aided drug design with the help of available data, especially the *in silico*-based quantitative structure-activity relationship (QSAR) modelling techniques have emerged as a promising alternative tool toward the effective screening of potential drugs. Also, structure, reactivity, and interaction of anticancer drugs with DNA are very crucial in understanding their anticancer activities. These properties can only be properly described when the laws of quantum mechanics are taken into account. In the present thesis we have used the most popular DFT method to perform structure-reactivity studies of platinum based

and organic anticancer drug molecules in gas and solvent phases with the help of DMol^3 program. We have investigated the binding mechanism of platinum drug using Gaussian 03 program, whereas QM/MM based ONIOM method has been used to study its interaction with DNA.

Influence of basis sets on electron affinities (EAs) of the DNA- and RNA bases has been investigated using B3LYP functional with different basis sets (6-31G, TZVP, 6-311++G**). Then we studied the effect of some PBE functionals namely, PBEOP, PBELYP, and PBEVWN on EA values of the nucleobases using basis set which predicted the most reliable values with B3LYP functional. Observation of the trends in the values of EA and dipole moment of the molecules enable us to identify the features of a basis set that shows the presence of dipole-bound state of some of the nucleobases. The vertical electron affinities with B3LYP and PBEOP functionals are close to the experimental values. Our calculated adiabatic electron affinity values of uracil and thymine are positive for basis set with diffuse functions using B3LYP functional. On the other hand, for adenine, guanine, and cytosine we found unstable anion at all levels of basis sets and functionals.

We have calculated the DFT based reactivity descriptors of some *cis*-platinum(II) complexes, including clinically used drug molecules, cisplatin, carboplatin, and oxaliplatin in gas and solvent phases to investigate their reactive nature. We observed that solvent phase predicted global reactivity trend of the drugs correlates well with the available experimental results. From the calculations of local reactivity descriptors, we found Pt sites are prone to nucleophilic attack with maximum values of f^+ and ω^+. Simple regression analysis is applied to build up a QSAR model based on DFT derived electrophilicity index for the complexes against A2780 human ovarian adenocarcinoma cell line to establish the importance of the descriptor in predicting anticancer activities.

Cytotoxic activities of several *cis*-platinum complexes against A2780 human ovarian adenocarcinoma cell line and its cisplatin resistant subline (A2780Cp8) have been investigated by QSAR analysis using DFT based descriptors. It has been observed that three parameters i.e., ω, ω^+, and E_{NL} provides regression models capable of predicting $\log(IC_{50}^{-1})$ of the complexes in both gas and solvent phases. Within the study we have also assessed hydrophobicity ($\log P_{o/w}$) of a set of 24 *cis*-

platinum complexes using QSPR approach. Our established QSPR equations modelled by ω, ω^+, MR_{CL}, and SA parameters against five different concentrations of MeOH (0%–50%) show a very good statistical quality. The $\log P_{o/w}$ values of an additional set of 20 *cis*-platinum complexes (training set) calculated using the same descriptors are in close proximity to their experimental values. Then we have investigated the predictive power of the the model by calculating $\log P_{o/w}$ of 4 compounds in the test set and found their predicted values to be in good agreement with the experimental ones. We have observed that for each QSPR equation, solvent model played an important role by increasing the internal predictive ability of the model.

Next we perform QSAR analysis of some nucleoside analogues by using different quantum-mechanical parameters including DFT based reactivity descriptors. This study also supports the application of reactivity descriptors in QSAR analysis. Our calculated equations for a set of 14 carbocyclic nucleosides against three cancer cell lines in gas and solvent phases suggest that lower values of E_{NL} combined with higher values of ω and SA, increases inhibitory activities. Based on this observation we have designed 10 new compounds with rather high anticancer activities. QSAR models for another set of 20 nucleoside analogues with three parameters i.e., ω, E_{NL}, and SA provide significant statistical parameters suggesting the usefulness of the selected descriptors in the prediction of anticancer activity of nucleosides.

In this thesis, we have carried out a systematic investigation to elucidate the binding mechanism of cisplatin analogue *cis*-[$PtCl_2(NH_3)$(2-picoline)] (AMD473). While producing marked activity, this orally active drug has an acceptable safety profile with less toxicity. Out of four different pathways of hydrolysis, we found the rate of hydrolysis of chloride ligand *trans* to 2-picoline group is higher than that of chloride ligand *cis* to 2-picoline group. We observed that the activation barrier for second step of hydrolysis is higher than the first step. This indicates that monoaquo forms are the most probable species that react with DNA in agreement with the experimental observation. Further monofunctional binding of the hydrated species with guanine is also studied to obtain detailed insight into binding mechanism of the drug. From the calculated activation barriers for three guanine substitution reactions

in gas and solvent phases we have suggested a most probable path for guanine binding.

To study the stability of AMD473-DNA adduct, we have performed ONIOM calculations of four different forms of the adduct. It is observed that the stereo-isomer with 2-picoline *trans* to 5'-G and 2-methyl group directed towards the DNA major groove center is the most stable structure from both energetical and structural point of views. Then we investigated the intermolecular proton transfer between the base pairs of the most stable structure. We found that the single proton transfer between N1 position of guanine and N3 position of cytosine gives an energetically stable structure while no simultaneous single proton transfer is energetically possible. The small difference in energy between the protonated products and the original one explains the damage of DNA when the drug molecule binds with it.

Future Scopes

Several lines of research would be interesting to pursue—

1) The application of DFT based global and local reactivity descriptors can be extended in the prediction of biological activity of different drug molecules.

2) Studies can be made to investigate the effect of different carrier and leaving ligands on the reactivity of platinum drugs.

3) Based on the developed QSAR models, new drug molecules with higher activities can be designed.

4) Studies on binding mechanism of platinum drugs and the properties of their adduct with DNA will be useful to enhance the activity and reduce the toxicity of drug molecules.

LIST OF PUBLICATIONS

In Journals

1) **P. Sarmah** and R. C. Deka, *Stability of AMD473(cis-[$PtCl_2(NH_3)$(2picoline)]) Drug-DNA Adduct and Proton Transfer in DNA Base Pairs : A Theoretical Study* (communicated).

2) N. Barua, **P. Sarmah,** R. C. Deka and A. K. Buragohain, *DFT-based QSAR Models to Predict the Antimycobacterial Activity of Chalcones* (communicated).

3) K. K. Hazarika, **P. Sarmah,** N. C. Barua and R. C. Deka, *Density Functional Theory Studies on Reactivity of Artemisinin and Some of its Derivatives* (communicated).

4) **P. Sarmah** and R. C. Deka, *Hydrolysis and Binding Mechanism of AMD473 (cis-[$PtCl_2(NH_3)$(2-picoline)]) with Guanine: A Quantum Mechanical Study,* **J. Mol. Struct. THEOCHEM** 955, (2010) 53.

5) **P. Sarmah** and R. C. Deka, *Anticancer Activity of Nucleoside Analogues: A Density Functional Theory Based QSAR Study*, **J. Mol. Model.** 16 (2010) 411.

6) **P. Sarmah** and R. C. Deka, *DFT-based QSAR and QSPR Models of Several cis-Platinum Complexes: Solvent Effect,* **J. Compt. Aided .Mol. Des.** 23 (2009) 343.

7) **P. Sarmah** and R. C. Deka, *Solvent Effect on the Reactivity of cis-Platinum (II) complexes: A Density Functional Approach*, **Int. J. Quant. Chem.** 108 (2008) 1400.

8) **P. Sarmah** and R. C. Deka, *Density Functional Studies on the Electron Affinities of DNA and RNA Bases*, **Mol. Simulation** 34 (2008) 879.

In Conferences

1) **P. Sarmah** and R. C. Deka, *Density Functional Studies on Structure-Activity Relationship and Binding Mechanism of Platinum Anticancer Drugs,* **Recent Developments & Trends in Computational Chemistry,** NEHU, Shillong, March 12-13, 2010.

2) **P. Sarmah** and R. C. Deka, *DFT-Based Reactivity and QSAR Studies on Anticancer Activity of cis-Platinum (II) Complexes*, **12th ISCBC International Sysposium, BITS,** Pillani, Rajasthan, February 22-24, 2008.

3) **P. Sarmah** and R. C. Deka, *Density Functional Based Global Reactivity of Platinum Anticancer Drugs,* **9th CRSI National Symposium in Chemistry,** Department of Chemistry, Delhi University, Delhi, Frebruary 1-4, 2007.

4) **P. Sarmah** and R. C. Deka, *Fragment Molecular Orbital Method: A Computational Method for Large Molecules*, **National Symposium on Condensed Matter Days**, Tezpur University, August 29-31, 2006.